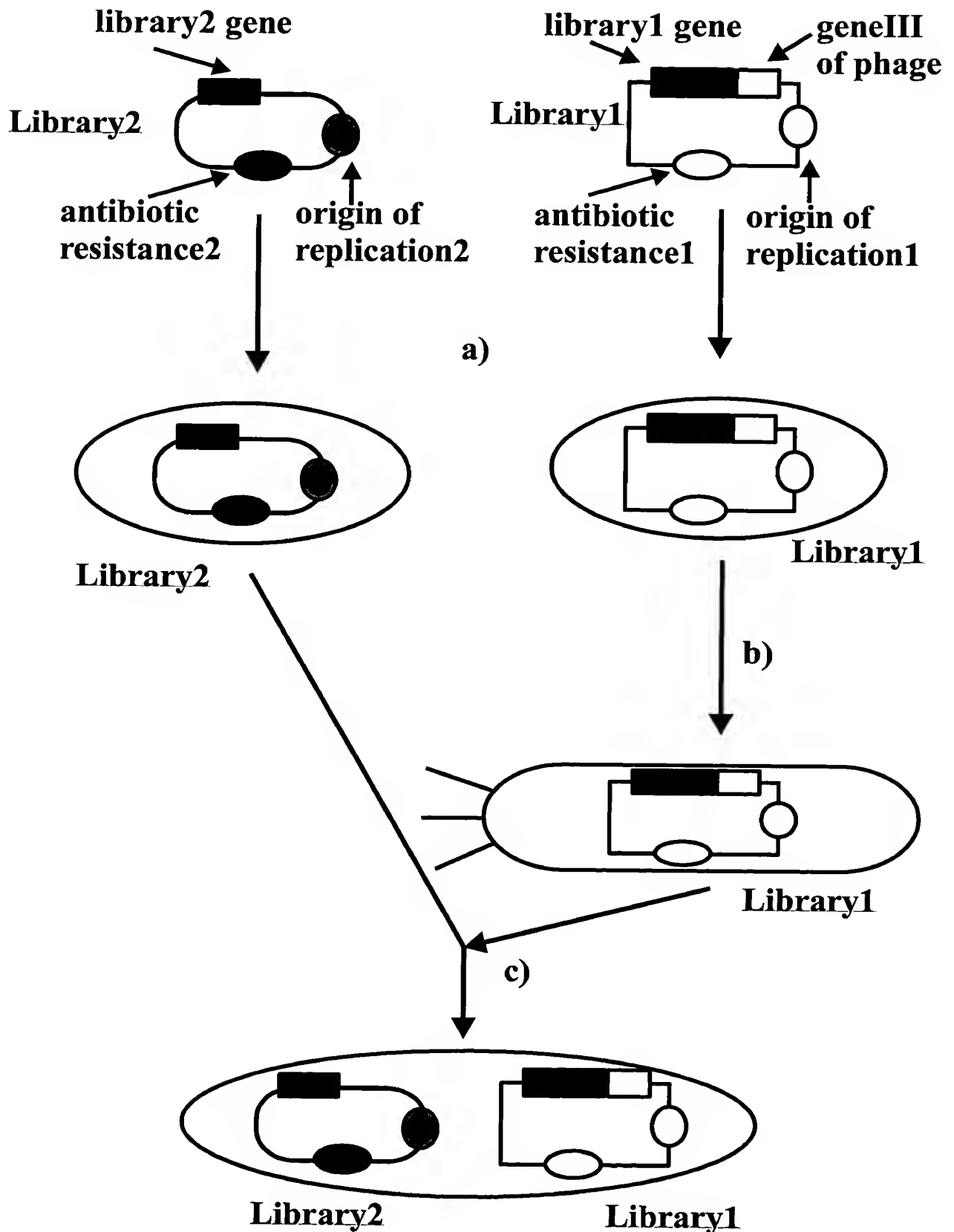


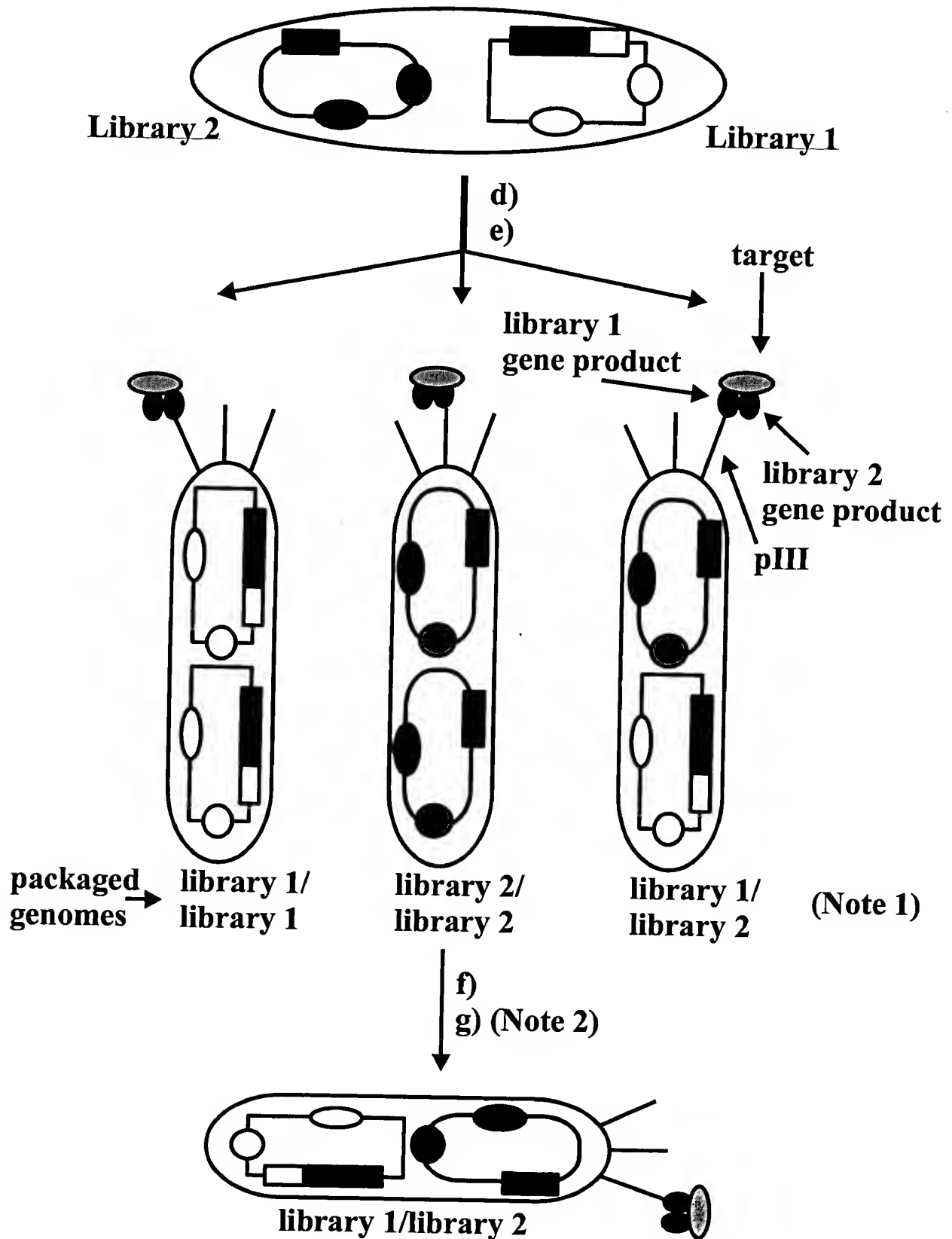
1/39

Figure 1: General description of the polyphage principle



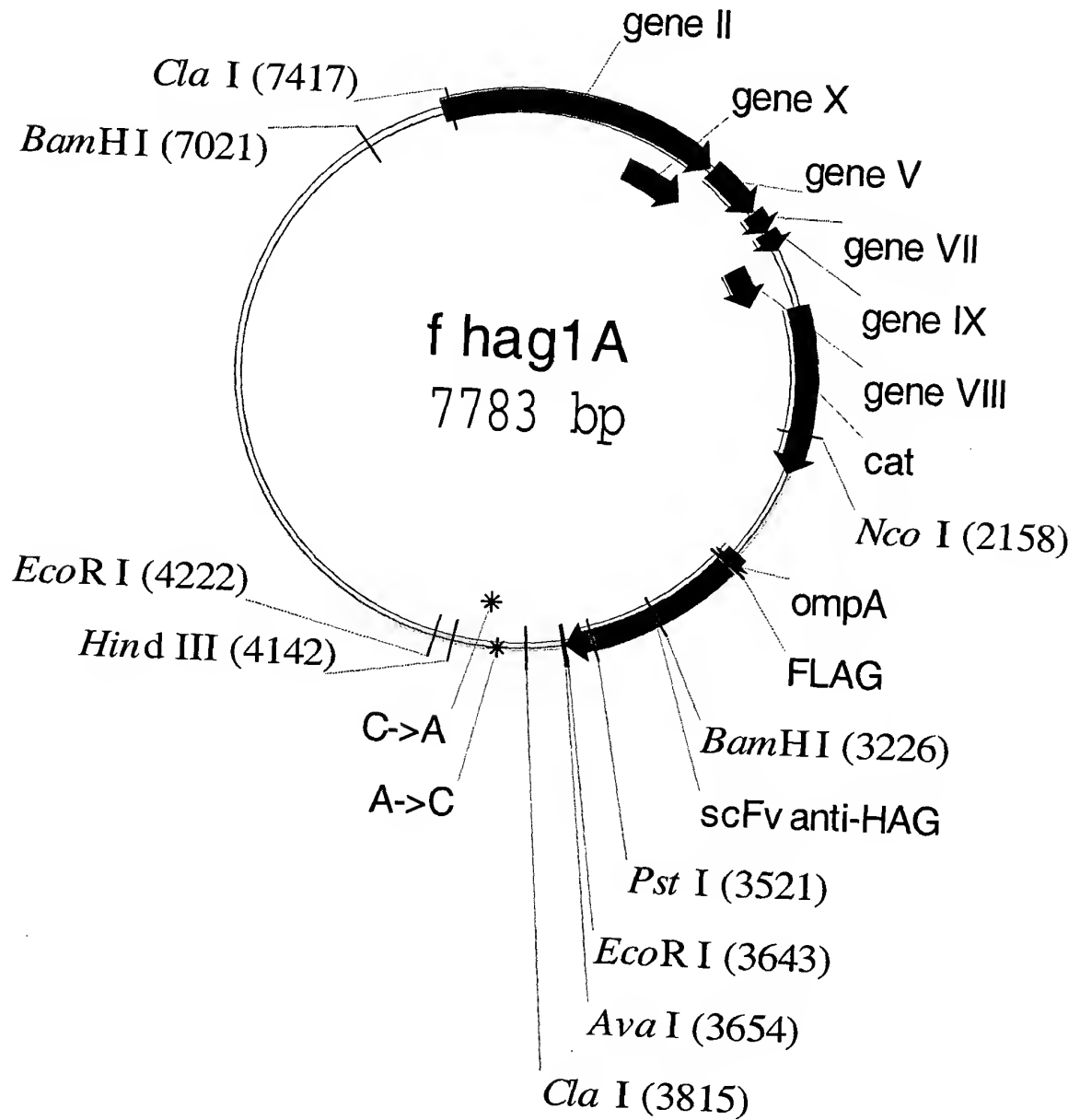
2/39

Figure 1: General description of the polyphage principle (cont.)



3/39

Figure 2



4/39

1	AACGCTACTA	CCATTAGTAG	AATTGATGCC	ACCTTTTTCAG	CTCGCGCCCC
	TTGCGATGAT	GGTAATCATC	TTAACTACGG	TGGAAAAGTC	GAGCGCGGGG
51	AAATGAAAAT	ATAGCTAAAC	AGGTTATTGA	CCATTTGCGA	AATGTATCTA
	TTTACTTTTA	TATCGATTTG	TCCAATAACT	GGTAAACGCT	TTACATAGAT
101	ATGGTCAAAC	TAAATCTACT	CGTTCGCAGA	ATTGGGAATC	AACTGTTACA
	TACCAGTTTG	ATTTAGATGA	GCAAGCGTCT	TAACCCTTAG	TTGACAATGT
151	TGGAATGAAA	CTTCCAGACA	CCGTACTTTA	GTTGCATATT	TAAAACATGT
	ACCTTACTTT	GAAGGTCTGT	GGCATGAAAT	CAACGTATAA	ATTTTGTACA
201	TGAACTACAG	CACCAGATTC	AGCAATTAAG	CTCTAAGCCA	TCCGCAAAAA
	ACTTGATGTC	GTGGTCTAAG	TCGTTAATTC	GAGATTCGGT	AGGCGTTTTT
251	TGACCTCTTA	TCAAAAAGGAG	CAATTAAAGG	TACTGTCTAA	TCCTGACCTG
	ACTGGAGAAT	AGTTTTCCCTC	GTTAATTTCC	ATGACAGATT	AGGACTGGAC
301	TTGGAATTTG	CTTCCGGTCT	GGTTCGCTTT	GAGGCTCGAA	TTGAAACGCG
	AACCTTAAAC	GAAGGCCAGA	CCAAGCGAAA	CTCCGAGCTT	AACTTTGCGC
351	ATATTTGAAG	TCTTTCGGGC	TTCTCTTTAA	TCTTTTTGAT	GCAATTCGCT
	TATAAACTTC	AGAAAGCCCG	AAGGAGAATT	AGAAAACTA	CGTTAAGCGA
401	TTGCTTCTGA	CTATAATAGA	CAGGGTAAAG	ACCTGATTTT	TGATTTATGG
	AACGAAGACT	GATATTATCT	GTCCCATTTC	TGGACTAAAA	ACTAAATACC
451	TCATTCTCGT	TTTCTGAACT	GTTTAAAGCA	TTTGAGGGGG	ATTCAATGAA
	AGTAAGAGCA	AAAGACTTGA	CAAATTTTCGT	AAACTCCCCC	TAAGTTACTT
501	TATTTATGAC	GATTCCGCAG	TATTGGACGC	TATCCAGTCT	AAACATTTTA
	ATAAATACTG	CTAAGGCGTC	ATAACCTGCG	ATAGGTCAGA	TTTGTAATAA
551	CAATTACCCC	CTCTGGCAAA	ACTTCCTTTG	CAAAAGCCTC	TCGCTATTTT
	GTTAATGGGG	GAGACCGTTT	TGAAGGAAAC	GTTTTTCGGAG	AGCGATAAAA
601	GGTTTCTATC	GTCGTCTGGT	TAATGAGGGT	TATGATAGTG	TTGCTCTTAC
	CCAAAGATAG	CAGCAGACCA	ATTACTCCCA	ATACTATCAC	AACGAGAATG
651	CATGCCTCGT	AATTCCTTTT	GGCGTTATGT	ATCTGCATTA	GTTGAGTGTG
	GTACGGAGCA	TTAAGGAAAA	CCGCAATACA	TAGACGTAAT	CAACTCACAC
701	GTATTCCTAA	ATCTCAATTG	ATGAATCTTT	CCACCTGTAA	TAATGTTGTT
	CATAAGGATT	TAGAGTTAAC	TACTTAGAAA	GGTGGACATT	ATTACAACAA
751	CCGTTAGTTC	GTTTTATTAA	CGTAGATTTT	TCCTCCCAAC	GTCCTGACTG
	GGCAATCAAG	CAAAATAATT	GCATCTAAAA	AGGAGGGTTG	CAGGACTGAC
801	GTATAATGAG	CCAGTTCTTA	AAATCGCATA	AGGTAATTCA	AAATGATTAA
	CATATTACTC	GGTCAAGAAT	TTTAGCGTAT	TCCATTAAAGT	TTTACTAATT
851	AGTTGAAATT	AAACCGTCTC	AAGCGCAATT	TACTACCCGT	TCTGGTGTTT
	TCAACTTTAA	TTTGGCAGAG	TTGCGGTTAA	ATGATGGGCA	AGACCACAAA

5/39

901	CTCGTCAGGG	CAAGCCTTAT	TCACTGAATG	AGCAGCTTTG	TTACGTTGAT
	GAGCAGTCCC	GTTTCGGAATA	AGTGACTTAC	TCGTTCGAAAC	AATGCAACTA
951	TTGGGTAATG	AATATCCGGT	GCTTGTCAAG	ATTACTCTCG	ACGAAGGTCA
	AACCCATTAC	TTATAGGCCA	CGAACAGTTC	TAATGAGAGC	TGCTTCCAGT
1001	GCCAGCGTAT	GCGCCTGGTC	TGTACACCGT	GCATCTGTCC	TCGTTCAAAG
	CGGTTCGCATA	CGCGGACCAG	ACATGTGGCA	CGTAGACAGG	AGCAAGTTTC
1051	TTGGTCAGTT	CGGTTCTCTT	ATGATTGACC	GTCTGCGCCT	CGTTCCGGCT
	AACCAGTCAA	GCCAAGAGAA	TACTAACTGG	CAGACGCGGA	GCAAGGCCGA
1101	AAGTAACATG	GAGCAGGTCG	CGGATTTCTGA	CACAATTTAT	CAGGCGATGA
	TTCATTGTAC	CTCGTCCAGC	GCCTAAAGCT	GTGTTAAATA	GTCCGCTACT
1151	TACAAATCTC	CGTTGTACTT	TGTTTTCGCGC	TTGGTATAAT	CGCTGGGGGT
	ATGTTTAGAG	GCAACATGAA	ACAAAGCGCG	AACCATATTA	GCGACCCCCA
1201	CAAAGATGAG	TGTTTTAGTG	TATTCTTTTCG	CCTCTTTTCGT	TTTAGGTTGG
	GTTTCTACTC	ACAAAATCAC	ATAAGAAAGC	GGAGAAAGCA	AAATCCAACC
1251	TGCCTTCGTA	GTGGCATTAC	GTATTTTACC	CGTTTAATGG	AAACTTCCTC
	ACGGAAGCAT	CACCGTAATG	CATAAAATGG	GCAAATTACC	TTTGAAGGAG
1301	ATGCGTAAGT	CTTTAGTCCT	CAAAGCCTCC	GTAGCCGTTG	CTACCCTCGT
	TACGCATTCA	GAAATCAGGA	GTTTCGGAGG	CATCGGCAAC	GATGGGAGCA
1351	TCCGATGCTG	TCTTTCGCTG	CTGAGGGTGA	CGATCCCGCA	AAAGCGGCCT
	AGGCTACGAC	AGAAAGCGAC	GACTCCCACT	GCTAGGGCGT	TTTCGCCGGA
1401	TTGACTCCCT	GCAAGCCTCA	GCGACCGAAT	ATATCGGTTA	TGCGTGGGCG
	AACTGAGGGA	CGTTCGGAGT	CGCTGGCTTA	TATAGCCAAT	ACGCACCCGC
1451	ATGGTTGTTG	TCATTGTCTG	CGCAACTATC	GGTATCAAGC	TGTTTAAGAA
	TACCAACAAC	AGTAACAGCC	GCGTTGATAG	CCATAGTTCTG	ACAAATTCTT
1501	ATTCACCTCG	AAAGCAAGCT	GATAAAGGAG	GTTTCTCGAT	CGAGACGTTN
	TAAGTGGAGC	TTTCGTTCGA	CTATTTCTCT	CAAAGAGCTA	GCTCTGCAAN
1551	NNNGAGGTTC	CAACTTTCAC	CATAATGAAA	TAAGATCACT	ACCGGGCGTA
	NNNCTCCAAG	GTTGAAAGTG	GTATTACTTT	ATTCTAGTGA	TGGCCCGCAT
1601	TTTTTTGAGT	TATCGAGATT	TTCAGGAGCT	AAGGAAGCTA	AAATGGAGAA
	AAAAAACTCA	ATAGCTCTAA	AAGTCCTCGA	TTCTTTCGAT	TTTACCTCTT
1651	AAAAATCACT	GGATATACCA	CCGTTGATAT	ATCCCAATGG	CATCGTAAAG
	TTTTTAGTGA	CCTATATGGT	GGCAACTATA	TAGGGTTACC	GTAGCATTTT
1701	AACATTTTGA	GGCATTTTCAG	TCAGTTGCTC	AATGTACCTA	TAACCAGACC
	TTGTAAAACT	CCGTAAAGTC	AGTCAACGAG	TTACATGGAT	ATTGGTCTGG
1751	GTTCAGCTGG	ATATTACGGC	CTTTTTTAAAG	ACCGTAAAGA	AAAATAAGCA
	CAAGTCGACC	TATAATGCCG	GAAAAATTTT	TGGCATTTCT	TTTTATTCGT

6/39

1801	CAAGTTTTAT	CCGGCCTTTA	TTCACATTCT	TGCCCCCCTG	ATGAATGCTC
	GTTCAAAATA	GGCCGGAAAT	AAGTGTAAGA	ACGGGCGGAC	TACTTACGAG
1851	ATCCGGAGTT	CCGTATGGCA	ATGAAAGACG	GTGAGCTGGT	GATATGGGAT
	TAGGCCTCAA	GGCATACCGT	TACTTTCTGC	CACTCGACCA	CTATACCCTA
1901	AGTGTTTACC	CTTGTTACAC	CGTTTTCCAT	GAGCAAACCTG	AAACGTTTTTC
	TCACAAGTGG	GAACAATGTG	GCAAAAGGTA	CTCGTTTGAC	TTTGCAAAAG
1951	ATCGCTCTGG	AGTGAATACC	ACGACGATTT	CCGGCAGTTT	CTACACATAT
	TAGCGAGACC	TCACTTATGG	TGCTGCTAAA	GGCCGTCAAA	GATGTGTATA
2001	ATTCGCAAGA	TGTGGCGTGT	TACGGTGAAA	ACCTGGCCTA	TTTCCCTAAA
	TAAGCGTTCT	ACACCGCACA	ATGCCACTTT	TGGACCGGAT	AAAGGGATTT
2051	GGGTTTATTG	AGAATATGTT	TTTCGTCTCA	GCCAATCCCT	GGGTGAGTTT
	CCCAAATAAC	TCTTATACAA	AAAGCAGAGT	CGGTTAGGGA	CCCACTCAAA
2101	CACCAGTTTT	GATTTAAACG	TGGCCAATAT	GGACAACCTC	TTCGCCCCCG
	GTGGTCAAAA	CTAAATTTGC	ACCGGTTATA	CCTGTTGAAG	AAGCGGGGGC

NcoI

~~~~~

|      |            |             |            |            |            |
|------|------------|-------------|------------|------------|------------|
| 2151 | TTTTTACCAT | GGGCAAATAT  | TATACGCAAG | GCGACAAGGT | GCTGATGCCG |
|      | AAAAGTGGTA | CCCGTTTATA  | ATATGCGTTC | CGCTGTTCCA | CGACTACGGC |
| 2201 | CTGGCGATTG | AGGTTTCATCA | TGCCGTCTGT | GATGGCTTCC | ATGTCGGCAG |
|      | GACCGCTAAG | TCCAAGTAGT  | ACGGCAGACA | CTACCGAAGG | TACAGCCGTC |
| 2251 | AATGCTTAAT | GAATTACAAC  | AGTACTGCGA | TGAGTGGCAG | GGCGGGGCGT |
|      | TTACGAATTA | CTTAATGTTG  | TCATGACGCT | ACTCACCGTC | CCGCCCCGCA |
| 2301 | AATTTTTTTA | AGGCAGTTAT  | TGGTGCCCTT | AAACGCCTGG | TGCTACGCCT |
|      | TTAAAAAAT  | TCCGTCAATA  | ACCACGGGAA | TTTGCGGACC | ACGATGCGGA |
| 2351 | GAATAAGTGA | TAATAAGCGG  | ATGAATGGCA | GAAATTCGAA | AGCAAATTCG |
|      | CTTATTTACT | ATTATTCGCC  | TACTTACCGT | CTTTAAGCTT | TCGTTTAAGC |
| 2401 | ACCCGGTCGT | CGGTTTCAGGG | CAGGGTCGTT | AAATAGCCGC | TTATGTCTAT |
|      | TGGGCCAGCA | GCCAAGTCCC  | GTCCCAGCAA | TTTATCGGCG | AATACAGATA |
| 2451 | TGCTGGTTTA | CCGGTTTATT  | GACTACCGGA | AGCAGTGTGA | CCGTGTGCTT |
|      | ACGACCAAAT | GGCCAAATAA  | CTGATGGCCT | TCGTCACACT | GGCACACGAA |
| 2501 | CTCAAATGCC | TGAGGCCAGT  | TTGCTCAGGC | TCTCCCCGTG | GAGGTAATAA |
|      | GAGTTTACGG | ACTCCGGTCA  | AACGAGTCCG | AGAGGGGCAC | CTCCATTATT |
| 2551 | TTGCTCGACC | GATAAAAGCG  | GCTTCCTGAC | AGGAGGCCGT | TTTGTTTTGC |
|      | AACGAGCTGG | CTATTTTCGC  | CGAAGGACTG | TCCTCCGGCA | AAACAAAACG |
| 2601 | AGCCACCTC  | AACGCAATTA  | ATGTGAGTTA | GCTCACTCAT | TAGGCACCCC |
|      | TCGGGTGGAG | TTGCGTTAAT  | TACACTCAAT | CGAGTGAGTA | ATCCGTGGGG |
| 2651 | AGGCTTTACA | CTTTATGCTT  | CCGGCTCGTA | TGTTGTGTGG | AATTGTGAGC |
|      | TCCGAAATGT | GAAATACGAA  | GGCCGAGCAT | ACAACACACC | TTAACACTCG |

7/39

|      |            |             |             |            |            |
|------|------------|-------------|-------------|------------|------------|
| 2701 | GGATAACAAT | TTCACACAGG  | AAACAGCTAT  | GACCATGATT | ACGAATTTCT |
|      | CCTATTGTTA | AAGTGTGTCC  | TTTGTGCGATA | CTGGTACTAA | TGCTTAAAGA |
| 2751 | AGATAACGAG | GGCAAATCAT  | GAAAAAGACA  | GCTATCGCGA | TTGCAGTGGC |
|      | TCTATTGCTC | CCGTTTAGTA  | CTTTTCTCTGT | CGATAGCGCT | AACGTCACCG |
| 2801 | ACTGGCTGGT | TTCGCTACCG  | TAGCGCAGGC  | CGACTACAAA | GATATCGTTA |
|      | TGACCGACCA | AAGCGATGGC  | ATCGCGTCCG  | GCTGATGTTT | CTATAGCAAT |
| 2851 | TGACCCAGTC | ACCGTCCTCC  | CTGACCGTTA  | CCGCTGGTGA | AAAAGTTACC |
|      | ACTGGGTCAG | TGGCAGGAGG  | GACTGGCAAT  | GGCGACCACT | TTTTCAATGG |
| 2901 | ATGTCCTGCA | CCTCCTCCCA  | GTCCCTGTTC  | AACTCCGGTA | AACAGAAAAA |
|      | TACAGGACGT | GGAGGAGGGT  | CAGGGACAAG  | TTGAGGCCAT | TTGTCTTTTT |
| 2951 | CTACCTGACC | TGGTATCAGC  | AGAAACCGGG  | TCAGCCACCG | AAAGTTCTGA |
|      | GATGGACTGG | ACCATAGTCG  | TCTTTGGCCC  | AGTCGGTGGC | TTTCAAGACT |
| 3001 | TCTACTGGGC | TTCCACCCGT  | GAATCCGGTG  | TTCCAGACCG | TTTCACCGGT |
|      | AGATGACCCG | AAGGTGGGCA  | CTTAGGCCAC  | AAGGTCTGGC | AAAGTGGCCA |
| 3051 | TCCGGTTCCG | GCACCGACTT  | CACCCTGACC  | ATCTCCTCCG | TTCAGGCTGA |
|      | AGGCCAAGGC | CGTGGCTGAA  | GTGGGACTGG  | TAGAGGAGGC | AAGTCCGACT |
| 3101 | AGACCTGGCT | GTTTACTACT  | GCCAGAACGA  | CTACTCCAAC | CCACTGACCT |
|      | TCTGGACCGA | CAAATGATGA  | CGGTCTTGCT  | GATGAGGTTG | GGTGACTGGA |
| 3151 | TCGGTGGTGG | CACCAAACCTG | GAACTTAAGC  | GCGCTGGTGG | TGGAGGGTCT |
|      | AGCCACCACC | GTGGTTTGAC  | CTTGAATTCG  | CGCGACCACC | ACCTCCCAGA |
|      |            |             | BamHI       |            |            |
|      |            |             | ~~~~~       |            |            |
| 3201 | GGAGGAGGTG | GGAGTGGGGG  | AGGTGGATCC  | GGCGGGGGAG | GTTCAGGGGG |
|      | CCTCCTCCAC | CCTCACCCCC  | TCCACCTAGG  | CCGCCCCCTC | CAAGTCCCCC |
| 3251 | TGGCGGTAGT | GGAGGGGGCG  | GTTCAGAAAGT | TCAACTAGTT | GAATCCGGTG |
|      | ACCGCCATCA | CCTCCCCCGC  | CAAGTCTTCA  | AGTTGATCAA | CTTAGGCCAC |
| 3301 | GTGACCTGGT | TAAACCGGGT  | GGTTCCTTGA  | AACTGTCCTG | CGCTGCTTCC |
|      | CACTGGACCA | ATTTGGCCCA  | CCAAGGGACT  | TTGACAGGAC | GCGACGAAGG |
| 3351 | GGTTTCTCCT | TCTCCTCCTA  | CGGTATGTCC  | TGGGTTCGTC | AGACCCCGGA |
|      | CCAAAGAGGA | AGAGGAGGAT  | GCCATACAGG  | ACCCAAGCAG | TCTGGGGCCT |
| 3401 | CAAACGTCTG | GAATGGGTTG  | CTACCATCTC  | CAACGGTGGT | GGTTACACCT |
|      | GTTTGCAGAC | CTTACCCAAC  | GATGGTAGAG  | GTTGCCACCA | CCAATGTGGA |
| 3451 | ACTACCCGGA | CTCCGTTAAA  | GGTCGTTTCA  | CCATCTCCCG | TGACAACGCT |
|      | TGATGGGCCT | GAGGCAATTT  | CCAGCAAAGT  | GGTAGAGGGC | ACTGTTGCGA |
|      |            |             | PstI        |            |            |
|      |            |             | ~~~~~       |            |            |
| 3501 | AAAAACACCC | TGTACCTGCA  | GATGTCCCTCC | CTGAAATCCG | AAGACTCAGC |

8/39

|      |            |            |              |            |             |
|------|------------|------------|--------------|------------|-------------|
|      | TTTTTGTGGG | ACATGGACGT | CTACAGGAGG   | GACTTTAGGC | TTCTGAGTCG  |
| 3551 | TATGTACTAC | TGCGCTCGTC | GTGAACGTTA   | CGACGAAAAC | GGTTTCGCTT  |
|      | ATACATGATG | ACGCGAGCAG | CACTTGCAAT   | GCTGCTTTTG | CCAAAGCGAA  |
|      |            |            |              | EcoRI      |             |
|      |            |            |              | ~~~~~      |             |
| 3601 | ACTGGGGTCA | GGGTACCCTG | GTTACCGTTT   | CAGCTTCCGG | AGAATTTCGAG |
|      | TGACCCAGT  | CCCATGGGAC | CAATGGCAAA   | GTCGAAGGCC | TCTTAAGCTC  |
|      | AvaI       |            |              |            |             |
|      | ~~~~~      |            |              |            |             |
| 3651 | GCCTCGGGGG | CCGAGGGCGG | CGGTTCTGGT   | TCCGGTGATT | TTGATTATGA  |
|      | CGGAGCCCCC | GGCTCCCGCC | GCCAAGACCA   | AGGCCACTAA | AACTAATACT  |
| 3701 | AAAAATGGCA | AACGCTAATA | AGGGGGCTAT   | GACCGAAAAT | GCCGATGAAA  |
|      | TTTTTACCGT | TTGCGATTAT | TCCCCGATA    | CTGGCTTTTA | CGGCTACTTT  |
| 3751 | ACGCGCTACA | GTCTGACGCT | AAAGGC AAAAC | TTGATTCTGT | CGCTACTGAT  |
|      | TGCGCGATGT | CAGACTGCGA | TTTCCGTTTG   | AACTAAGACA | GCGATGACTA  |
|      |            | Clai       |              |            |             |
|      |            | ~~~~~      |              |            |             |
| 3801 | TACGGTGCTG | CTATCGATGG | TTTCATTGGT   | GACGTTTCCG | GCCTTGCTAA  |
|      | ATGCCACGAC | GATAGCTACC | AAAGTAACCA   | CTGCAAAGGC | CGGAACGATT  |
| 3851 | TGGTAATGGT | GCTACTGGTG | ATTTTGCTGG   | CTCTAATTCC | CAAAATGGCTC |
|      | ACCATTACCA | CGATGACCAC | TAAAACGACC   | GAGATTAAGG | GTTTACCGAG  |
| 3901 | AAGTCGGTGA | CGGTGATAAT | TCACCTTTAA   | TGAATAATTT | CCGTCAATAT  |
|      | TTCAGCCACT | GCCACTATTA | AGTGGA AATT  | ACTTATTAAA | GGCAGTTATA  |
| 3951 | TTACCTTCCC | TCCCTCAATC | GGTTGAATGT   | CGCCCTTTTG | TCTTTGGCGC  |
|      | AATGGAAGGG | AGGGAGTTAG | CCAAC TTACA  | GCGGGAAAAC | AGAAACCGCG  |
| 4001 | TGGTAAACCA | TATGAATTTT | CTATTGATTG   | TGACAAAATA | AACTTATTCC  |
|      | ACCATTTGGT | ATACTTAAAA | GATAACTAAC   | ACTGTTTTAT | TTGAATAAGG  |
| 4051 | GTGGTGTCTT | TGCGTTTCTT | TTATATGTTG   | CCACCTTTAT | GTATGTATTT  |
|      | CACCACAGAA | ACGCAAAGAA | AATATAACAAC  | GGTGGAAATA | CATACATAAA  |
|      |            |            |              | HindIII    |             |
|      |            |            |              | ~~~~~      |             |
| 4101 | TCTACGTTTG | CTAACATACT | GCGTAATAAG   | GAGTCTTGAT | AAGCTTCGAG  |
|      | AGATGCAAAC | GATTGTATGA | CGCATTATTC   | CTCAGAACTA | TTCGAAGCTC  |
| 4151 | AAATTCACCT | CGAAAGCAAG | CTGATAAAACC  | GATACAATTA | AAGGCTCCTT  |
|      | TTTAAGTGGA | GCTTTCGTTT | GACTATTTGG   | CTATGTTAAT | TTCCGAGGAA  |
|      |            |            | EcoRI        |            |             |
|      |            |            | ~~~~~        |            |             |
| 4201 | TTGGAGCCTT | TTTTTTTGGG | GAATTCAATC   | ATGCCAGTTC | TTTTGGGTAT  |
|      | AACCTCGGAA | AAAAAAACCT | CTTAAGTTAG   | TACGGTCAAG | AAAACCCATA  |
| 4251 | TCCGTTATTA | TTGCGTTTCC | TCGGTTTCCT   | TCTGGTAACT | TTGTTCGGCT  |
|      | AGGCAATAAT | AACGCAAAGG | AGCCAAAGGA   | AGACCATTTA | AACAAGCCGA  |



9/39

|      |            |             |            |            |             |
|------|------------|-------------|------------|------------|-------------|
| 4301 | ATCTGCTTAC | TTTCCTTAAA  | AAGGGCTTCG | GTAAGATAGC | TATTGCTATT  |
|      | TAGACGAATG | AAAGGAATTT  | TTCCCGAAGC | CATTCTATCG | ATAACGATAA  |
| 4351 | TCATTGTTTC | TTGCTCTTAT  | TATTGGGCTT | AACTCAATTC | TTGTGGGTTA  |
|      | AGTAACAAAG | AACGAGAATA  | ATAACCCGAA | TTGAGTTAAG | AACACCCAAT  |
| 4401 | TCTCTCTGAT | ATTAGCGCAC  | AATTACCCTC | TGATTTTGT  | CAGGGCGTTC  |
|      | AGAGAGACTA | TAATCGCGTG  | TTAATGGGAG | ACTAAAACAA | GTCCCAGCAAG |
| 4451 | AGTTAATTCT | CCCGTCTAAT  | GCGCTTCCCT | GTTTTTATGT | TATTCTCTCT  |
|      | TCAATTAAGA | GGGCAGATTA  | CGCGAAGGGA | CAAAAATACA | ATAAGAGAGA  |
| 4501 | GTAAAGGCTG | CTATTTTTCAT | TTTTGACGTT | AAACAAAAAA | TCGTTTCTTA  |
|      | CATTTCCGAC | GATAAAAGTA  | AAAACGTCAA | TTTGTTTTTT | AGCAAAGAAT  |
| 4551 | TTTGGATTGG | GATAAATAAA  | TATGGCTGTT | TATTTTGTA  | CTGGCAAATT  |
|      | AAACCTAACC | CTATTTATTT  | ATACCGACAA | ATAAACATT  | GACCGTTTAA  |
| 4601 | AGGCTCTGGA | AAGACGCTCG  | TTAGCGTTGG | TAAGATTGAG | GATAAAATTG  |
|      | TCCGAGACCT | TTCTGCGAGC  | AATCGCAACC | ATTCTAAGTC | CTATTTTAAC  |
| 4651 | TAGCTGGGTG | CAAAATAGCA  | ACTAATCTTG | ATTTAAGGCT | TCAAAACCTC  |
|      | ATCGACCCAC | GTTTTATCGT  | TGATTAGAAC | TAAATTCCGA | AGTTTTGGAG  |
| 4701 | CCGCAAGTCG | GGAGGTTCGC  | TAAAACGCCT | CGCGTTCTTA | GAATACCGGA  |
|      | GGCGTTCAGC | CCTCCAAGCG  | ATTTTGCGGA | GCGCAAGAAT | CTTATGGCCT  |
| 4751 | TAAGCCTTCT | ATTTCTGATT  | TGCTTGCTAT | TGGTCGTGGT | AATGATTCCCT |
|      | ATTCGGAAGA | TAAAGACTAA  | ACGAACGATA | ACCAGCACCA | TTACTAAGGA  |
| 4801 | ACGACGAAAA | TAAAAACGGT  | TTGCTTGTTT | TTGATGAATG | CGGTACTTGG  |
|      | TGCTGCTTTT | ATTTTGTCCA  | AACGAACAAG | AACTACTTAC | GCCATGAACC  |
| 4851 | TTTAATACCC | GTTTATGGAA  | TGACAAGGAA | AGACAGCCGA | TTATTGATTG  |
|      | AAATTATGGG | CAAGTACCTT  | ACTGTTCCCT | TCTGTGCGCT | AATAACTAAC  |
| 4901 | GTTTCTTCAT | GCTCGTAAAT  | TGGGATGGGA | TATTATTTTT | CTTGTTTACG  |
|      | CAAAGAAGTA | CGAGCATTTA  | ACCCTACCCT | ATAATAAAAA | GAACAAGTCC  |
| 4951 | ATTTATCTAT | TGTTGATAAA  | CAGGCGCGTT | CTGCATTAGC | TGAACACGTT  |
|      | TAAATAGATA | ACAACATTTT  | GTCCGCGCAA | GACGTAATCG | ACTTGTGCAA  |
| 5001 | GTTTATTGTC | GCCGTCTGGA  | CAGAATTACT | TTACCCTTTG | TCGGCACTTT  |
|      | CAAATAACAG | CGGCAGACCT  | GTCTTAATGA | AATGGGAAAC | AGCCGTGAAA  |
| 5051 | ATATTCTCTT | GTTACTGGCT  | CAAAAATGCC | TCTGCCTAAA | TTACATGTTG  |
|      | TATAAGAGAA | CAATGACCGA  | GTTTTTACGG | AGACGGATTT | AATGTACAAC  |
| 5101 | GTGTTGTTAA | ATATGGTGAT  | TCTCAATTAA | GCCCTACTGT | TGAGCGTTGG  |
|      | CACAACAATT | TATACCACTA  | AGAGTTAATT | CGGGATGACA | ACTCGCAACC  |
| 5151 | CTTTATACTG | GTAAGAATTT  | ATATAACGCA | TATGACACTA | AACAGGCTTT  |
|      | GAAATATGAC | CATTCTTAAA  | TATATTGCGT | ATACTGTGAT | TTGTCCGAAA  |

10/39

|      |                          |                          |                          |                          |                           |
|------|--------------------------|--------------------------|--------------------------|--------------------------|---------------------------|
| 5201 | TTCCAGTAAT<br>AAGGTCATTA | TATGATTCAG<br>ATACTAAGTC | GTGTTTATTC<br>CACAAATAAG | ATATTTAACC<br>TATAAATTGG | CCTTATTTAT<br>GGAATAAATA  |
| 5251 | CACACGGTCG<br>GTGTGCCAGC | GTATTTCAAA<br>CATAAAGTTT | CCATTAAATT<br>GGTAATTTAA | TAGGTCAGAA<br>ATCCAGTCTT | GATGAAATTA<br>CTACTTTAAT  |
| 5301 | ACTAAAATAT<br>TGATTTTATA | ATTTGAAAAA<br>TAAACTTTTT | GTTTTCTCGC<br>CAAAGAGCG  | GTTCTTTGTC<br>CAAGAAACAG | TTGCGATAGG<br>AACGCTATCC  |
| 5351 | ATTTGCATCA<br>TAAACGTAGT | GCATTTACAT<br>CGTAAATGTA | ATAGTTATAT<br>TATCAATATA | AACCCAACCT<br>TTGGGTTGGA | AAGCCGGAGG<br>TTCGGCCTCC  |
| 5401 | TAAAAAAGGT<br>AATTTTTCCT | AGTCTCTCAG<br>TCAGAGAGTC | ACCTATGATT<br>TGGATACTAA | TTGATAAATT<br>AACTATTTAA | CACTATTGAC<br>GTGATAACTG  |
| 5451 | TCTTCTCAGC<br>AGAAGAGTCG | GTCTTAATCT<br>CAGAATTAGA | AAGCTATCGC<br>TTCGATAGCG | TATGTTTTCA<br>ATACAAAAGT | AGGATTCTAA<br>TCCTAAGATT  |
| 5501 | GGGAAAATTA<br>CCCTTTTAAT | ATTAATAGCG<br>TAATTATCGC | ACGATTTACA<br>TGCTAAATGT | GAAGCAAGGT<br>CTTCGTTCCA | TATTCCATCA<br>ATAAGGTAGT  |
| 5551 | CATATATTGA<br>GTATATAACT | TTTATGTACT<br>AAATACATGA | GTTTCAATTA<br>CAAAGTTAAT | AAAAAGGTAA<br>TTTTTCCATT | TTCAAATGAA<br>AAGTTTACTT  |
| 5601 | ATTGTTAAAT<br>TAACAATTTA | GTAATTAATT<br>CATTAATTAA | TTGTTTTCTT<br>AACAAAAGAA | GATGTTTGTT<br>CTACAAACAA | TCATCATCTT<br>AGTAGTAGAA  |
| 5651 | CTTTTGCTCA<br>GAAAACGAGT | AGTAATTGAA<br>TCATTAACTT | ATGAATAATT<br>TACTTATTAA | CGCCTCTGCG<br>GCGGAGACGC | CGATTTTCGTG<br>GCTAAAGCAC |
| 5701 | ACTTGGTATT<br>TGAACCATAA | CAAAGCAAAC<br>GTTTCGTTTG | AGGTGAATCT<br>TCCACTTAGA | GTTATTGTCT<br>CAATAACAGA | CACCTGATGT<br>GTGGACTACA  |
| 5751 | TAAAGGTACA<br>ATTTCCATGT | GTGACTGTAT<br>CACTGACATA | ATTCCTCTGA<br>TAAGGAGACT | CGTTAAGCCT<br>GCAATTCGGA | GAAAATTTAC<br>CTTTTAAATG  |
| 5801 | GCAATTTCTT<br>CGTTAAAGAA | TATCTCTGTT<br>ATAGAGACAA | TTACGTGCTA<br>AATGCACGAT | ATAATTTTGA<br>TATTAAAACT | TATGGTTGGC<br>ATACCAACCG  |
| 5851 | TCAATTCCTT<br>AGTTAAGGAA | CCATAATTCA<br>GGTATTAAGT | GAAATATAAC<br>CTTTATATTG | CCAAATAGTC<br>GGTTTATCAG | AGGATTATAT<br>TCCTAATATA  |
| 5901 | TGATGAATTG<br>ACTACTTAAC | CCATCATCTG<br>GGTAGTAGAC | ATATTCAGGA<br>TATAAGTCCT | ATATGATGAT<br>TATACTACTA | AATTCCGCTC<br>TTAAGGCGAG  |
| 5951 | CTTCTGGTGG<br>GAAGACCACC | TTTCTTTGTT<br>AAAGAAACAA | CCGCAAAATG<br>GGCGTTTTAC | ATAATGTTAC<br>TATTACAATG | TCAAACATTT<br>AGTTTGTAAA  |
| 6001 | AAAATTAATA<br>TTTTAATTAT | ACGTTTCGCG<br>TGCAAGCGCG | AAAGGATTTA<br>TTTCCTAAAT | ATAAGGGTTG<br>TATTCCTAAC | TAGAATTGTT<br>ATCTTAACAA  |
| 6051 | TGTTAAATCT<br>ACAATTTAGA | AATACATCTA<br>TTATGTAGAT | AATCCTCAAA<br>TTAGGAGTTT | TGTATTATCT<br>ACATAATAGA | GTTGATGGTT<br>CAACTACCAA  |

11/39

|      |            |             |             |            |             |
|------|------------|-------------|-------------|------------|-------------|
| 6101 | CTAACTTATT | AGTAGTTAGC  | CCCCCTAAAG  | ATATTTTAGA | TAACCTTCCG  |
|      | GATTGAATAA | TCATCAATCG  | CGGGGATTTT  | TATAAAATCT | ATTGGAAGGC  |
| 6151 | CAATTTCTTT | CTACTGTTGA  | TTTGCCAACT  | GACCAGATAT | TGATTGAAGG  |
|      | GTAAAGAAA  | GATGACAACT  | AAACGGTTGA  | CTGGTCTATA | ACTAACTTCC  |
| 6201 | ATTAATTTTC | GAGGTTTCAGC | AAGGTGATGC  | TTTAGATTTT | TCCTTTGCTG  |
|      | TAAATAAAAG | CTCCAAGTCG  | TTCCACTACG  | AAATCTAAAA | AGGAAACGAC  |
| 6251 | CTGGCTCTCA | GCGCGGCACT  | GTTGCTGGTG  | GTGTTAATAC | TGACCGTCTA  |
|      | GACCGAGAGT | CGCGCCGTGA  | CAACGACCAC  | CACAATTATG | ACTGGCAGAT  |
| 6301 | ACCTCTGTTT | TATCTTCTGC  | GGGTGGTTTC  | TTCGGTATTT | TTAACGGCGA  |
|      | TGGAGACAAA | ATAGAAGACG  | CCCACCAAGC  | AAGCCATAAA | AATTGCCGCT  |
| 6351 | TGTTTTAGGG | CTATCAGTTC  | GCGCATTAAG  | GAATAATAGC | CATTCAAAAA  |
|      | ACAAAATCCC | GATAGTCAAG  | CGCGTAATTT  | CTGATTATCG | GTAAGTTTTT  |
| 6401 | TATTGTCTGT | GCCTCGTATT  | CTTACGCTTT  | CAGGTCAGAA | GGGTTCTATT  |
|      | ATAACAGACA | CGGAGCATAA  | GAATGCGAAA  | GTCCAGTCTT | CCCAAGATAA  |
| 6451 | TCTGTTGGCC | AGAATGTCCC  | TTTTATTACT  | GGTCGTGTAA | CTGGTGAATC  |
|      | AGACAACCGG | TCTTACAGGG  | AAAATAATGA  | CCAGCACATT | GACCACTTAG  |
| 6501 | TGCCAATGTA | AATAATCCAT  | TTTACGCGGT  | TGAGCGTCAA | AATGTTGGTA  |
|      | ACGGTTACAT | TTATTAGGTA  | AAGTCTGCCA  | ACTCGCAGTT | TTACAACCAT  |
| 6551 | TTTCTATGAG | TGTTTTTCCC  | GTTGCAATGG  | CTGGCGGTAA | TATTGTTTTA  |
|      | AAAGATACTC | ACAAAAAGGG  | CAACGTTACC  | GACCGCCATT | ATAACAAAAT  |
| 6601 | GATATAACCA | GTAAGGCCGA  | TAGTTTGAGT  | TCTTCTACTC | AGGCAAGTGA  |
|      | CTATATTGGT | CATTCCGGCT  | ATCAAACCTA  | AGAAGATGAG | TCCGTTCACT  |
| 6651 | TGTTATTACT | AATCAAAGAA  | GTATTGCGAC  | AACGGTTAAT | TTGCGTGATG  |
|      | ACAATAATGA | TTAGTTTCTT  | CATAACGCTG  | TTGCCAATTA | AACGCACTAC  |
| 6701 | GTCAGACTCT | TTTGCTCGGT  | GGCCTCACTG  | ATTACAAAAA | CACCTTCTCAA |
|      | CAGTCTGAGA | AAACGAGCCA  | CCGAGGTGAC  | TAATGTTTTT | GTGAAGAGTT  |
| 6751 | GATTCTGGTG | TGCCGTTCCCT | GTCTAAAATC  | CCTTTAATCG | GCCTCCTGTT  |
|      | CTAAGACCAC | ACGGCAAGGA  | CAGATTTTAG  | GGAAATTAGC | CGGAGGACAA  |
| 6801 | TAGCTCCCGT | TCTGATTCTA  | ACGAGGAAAAG | CACGTTGTAC | GTGCTCGTCA  |
|      | ATCGAGGGCA | AGACTAAGAT  | TGCTCCTTTC  | GTGCAACATG | CACGAGCAGT  |
| 6851 | AAGCAACCAT | AGTACGCGCC  | CTGTAGCGGC  | GCATTAAGCG | CGGCGGGTGT  |
|      | TTCGTTGGTA | TCATGCGCGG  | GACATCGCCG  | CGTAATTTCG | GCCGCCACAA  |
| 6901 | GGTGGTTACG | CGCAGCGTGA  | CCGCTACACT  | TGCCAGCGCC | CTAGCGCCCC  |
|      | CCACCAATGC | GCGTCGCACT  | GGCGATGTGA  | ACGGTCGCGG | GATCGCGGGC  |
| 6951 | CTCCTTTTCG | TTTCTTCCCT  | TCCTTTCTCG  | CCACGTTCTC | CGGCTTTCCC  |
|      | GAGGAAAGCG | AAAGAAGGGA  | AGGAAAGAGC  | GGTGCAAGAG | GCCGAAAGGG  |

12/39

BamHI

~~~~~

7001	CGTCAAGCTC	TAAATCGGGG	GATCCCTTTA	GGGTTCGAT	TTAGTGCTTT
	GCAGTTCGAG	ATTTAGCCCC	CTAGGGAAAT	CCCAAGGCTA	AATCACGAAA
7051	ACGGCACCTC	GACCTCCAAA	AACTTGATTT	GGGTGATGGT	TCACGTAGTG
	TGCCGTGGAG	CTGGAGGTTT	TTGAACTAAA	CCCCTACCA	AGTGCATCAC
7101	GGCCATCGCC	CTGATAGACG	GTTTTTCGCC	CTTTGACGTT	GGAGTCCACG
	CCGGTAGCGG	GACTATCTGC	CAAAAAGCGG	GAAACTGCAA	CCTCAGGTGC
7151	TTCTTTAATA	GTGGACTCTT	GTTCCAAACT	GGAACAACAC	TCACAACATA
	AAGAAATTAT	CACCTGAGAA	CAAGGTTTGA	CCTTGTTGTG	AGTGTTGATT
7201	CTCGGCCTAT	TCTTTTGATT	TATAAGGATT	TTTGTCATTT	TCTGCTTACT
	GAGCCGGATA	AGAAAATAA	ATATTCCTAA	AAACAGTAAA	AGACGAATGA
7251	GGTTAAAAAA	TAAGCTGATT	TAACAAATAT	TTAACGCGAA	ATTTAACAAA
	CCAATTTTTT	ATTGACTTAA	ATTGTTTATA	AATTGCGCTT	TAAATTGTTT
7301	ACATTAACGT	TTACAATTTA	AATATTTGCT	TATACAATCA	TCCTGTTTTT
	TGTAATTGCA	AATGTTAAAT	TTATAAACGA	ATATGTTAGT	AGGACAAAAA
7351	GGGGCTTTTC	TGATTATCAA	CCGGGGTACA	TATGATTGAC	ATGCTAGTTT
	CCCCGAAAAG	ACTAATAGTT	GGCCCCATGT	ATACTAACTG	TACGATCAAA

ClaI

~~~~~

|      |            |             |            |            |            |
|------|------------|-------------|------------|------------|------------|
| 7401 | TACGATTACC | GTTTCATCGAT | TCTCTTGTTT | GCTCCAGACT | TTCAGGTAAT |
|      | ATGCTAATGG | CAAGTAGCTA  | AGAGAACAAA | CGAGGTCTGA | AAGTCCATTA |
| 7451 | GACCTGATAG | CCTTTGTAGA  | CCTCTCAAAA | ATAGCTACCC | TCTCCGGCAT |
|      | CTGGACTATC | GGAAACATCT  | GGAGAGTTTT | TATCGATGGG | AGAGGCCGTA |
| 7501 | GAATTTATCA | GCTAGAACGG  | TTGAATATCA | TATTGACGGT | GATTTGACTG |
|      | CTTAAATAGT | CGATCTTGCC  | AACTTATAGT | ATAACTGCCA | CTAAACTGAC |
| 7551 | TCTCCGGCCT | TTCTCACCCG  | TTTGAATCTT | TGCCTACTCA | TTACTCCGGC |
|      | AGAGGCCGGA | AAGAGTGGGC  | AAACTTAGAA | ACGGATGAGT | AATGAGGCCG |
| 7601 | ATTGCATTTA | AAATATATGA  | GGGTTCATAA | AATTTTTATC | CCTGCGTTGA |
|      | TAACGTAAAT | TTTATATACT  | CCCAAGATTT | TTAAAAATAG | GGACGCAACT |
| 7651 | AATTAAGGCT | TCACCAGCAA  | AAGTATTACA | GGGTCATAAT | GTTTTTGGTA |
|      | TTAATTCCGA | AGTGGTCGTT  | TTTATAATGT | CCCAGTATTA | CAAAAACCAT |
| 7701 | CAACCGATTT | AGCTTTATGC  | TCTGAGGCTT | TATTGCTTAA | TTTTGCTAAC |
|      | GTTGGCTAAA | TCGAAATACG  | AGACTCCGAA | ATAACGAATT | AAAACGATTG |
| 7751 | TCTCTGCCTT | GCTTGATACG  | TTTATTGGAT | GTT        |            |
|      | AGAGACGGAA | CGAACATGCT  | AAATAACCTA | CAA        |            |

Figure 2K

13/39

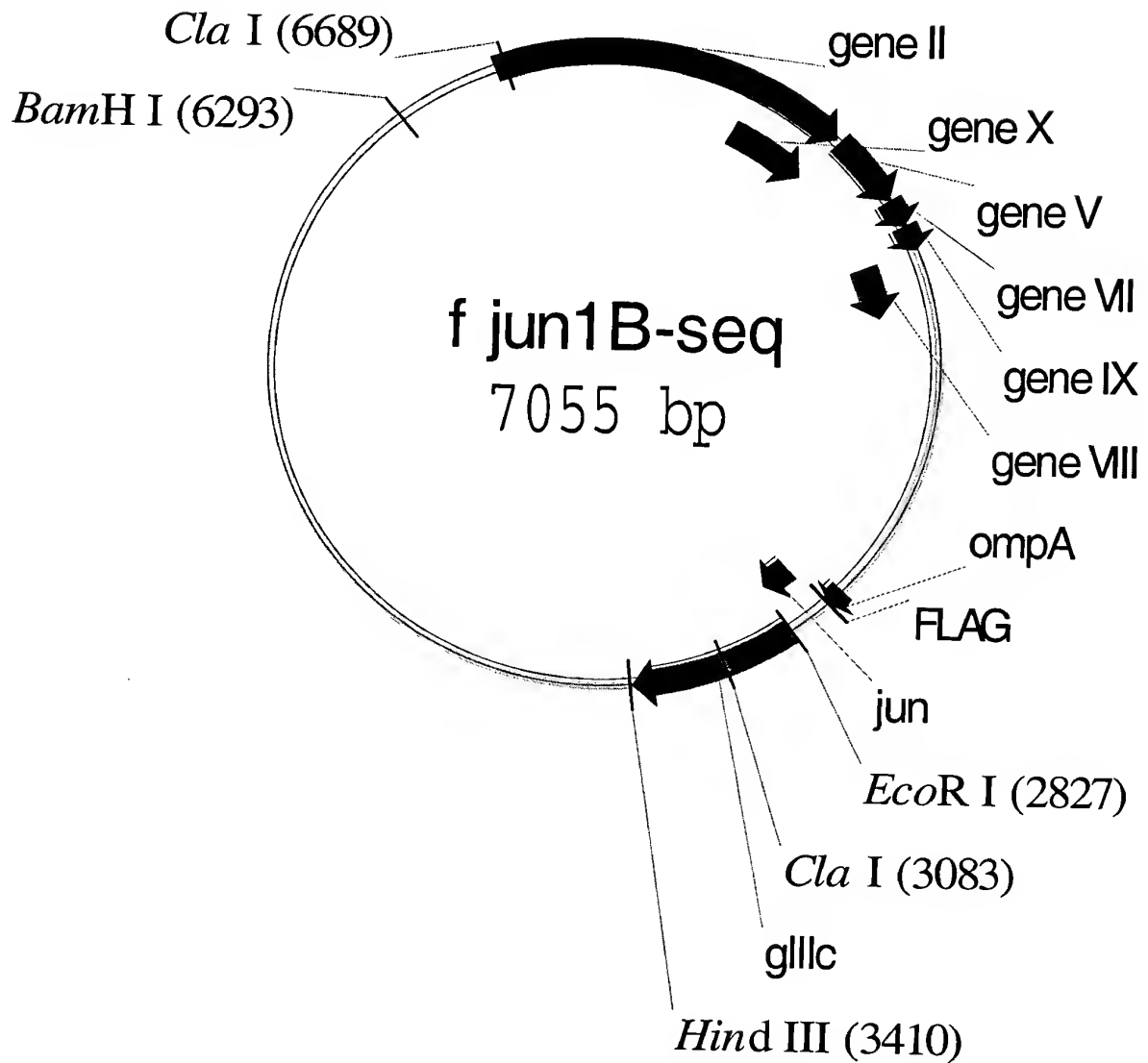
clai

~~~~~

7401	TACGATTACC	GTTTCATCGAT	TCTCTTGTTT	GCTCCAGACT	TTCAGGTAAT
	ATGCTAATGG	CAAGTAGCTA	AGAGAACAAA	CGAGGTCTGA	AAGTCCATTA
7451	GACCTGATAG	CCTTTGTAGA	CCTCTCAAAA	ATAGCTACCC	TCTCCGGCAT
	CTGGACTATC	GGAAACATCT	GGAGAGTTTT	TATCGATGGG	AGAGGCCGTA
7501	GAATTTATCA	GCTAGAACGG	TTGAATATCA	TATTGACGGT	GATTTGACTG
	CTTAAATAGT	CGATCTTGCC	AACCTATAGT	ATAACTGCCA	CTAAACTGAC
7551	TCTCCGGCCT	TTCTCACCCG	TTTGAATCTT	TGCCTACTCA	TTACTCCGGC
	AGAGGCCGGA	AAGAGTGGGC	AAACTTAGAA	ACGGATGAGT	AATGAGGCCG
7601	ATTGCATTTA	AAATATATGA	GGGTTCTAAA	AATTTTTATC	CCTGCGTTGA
	TAACGTAAAT	TTTATATACT	CCCAAGATTT	TTAAAAATAG	GGACGCAACT
7651	AATTAAGGCT	TCACCAGCAA	AAGTATTACA	GGGTCATAAT	GTTTTTGGTA
	TTAATTCCGA	AGTGGTCGTT	TTCATAATGT	CCCAGTATTA	CAAAAACCAT
7701	CAACCGATTT	AGCTTTATGC	TCTGAGGCTT	TATTGCTTAA	TTTGCTAAC
	GTTGGCTAAA	TCGAAATACG	AGACTCCGAA	ATAACGAATT	AAAACGATTG
7751	TCTCTGCCTT	GCTTGTACGA	TTTATTGGAT	GTT	
	AGAGACGGAA	CGAACATGCT	AAATAACCTA	CAA	

14/39

Figure 3



15/39

1	AACGCTACTA	CCATTAGTAG	AATTGATGCC	ACCTTTTTCAG	CTCGCGCCCC
	TTGCGATGAT	GGTAATCATC	TAACTACGG	TGGAAAAGTC	GAGCGCGGGG
51	AAATGAAAAT	ATAGCTAAAC	AGGTTATTGA	CCATTTGCGA	AATGTATCTA
	TTTACTTTTA	TATCGATTTG	TCCAATAACT	GGTAAACGCT	TTACATAGAT
101	ATGGTCAAAC	TAAATCTACT	CGTTCGCAGA	ATTGGGAATC	AACTGTTACA
	TACCAGTTTG	ATTTAGATGA	GCAAGCGTCT	TAACCCTTAG	TTGACAATGT
151	TGGAATGAAA	CTTCCAGACA	CCGTACTTTA	GTTGCATATT	TAAAACATGT
	ACCTTACTTT	GAAGGTCTGT	GGCATGAAAT	CAACGTATAA	ATTTTGTACA
201	TGAACTACAG	CACCAGATTC	AGCAATTAAG	CTCTAAGCCA	TCCGCAAAAA
	ACTTGATGTC	GTGGTCTAAG	TCGTTAATTC	GAGATTCGGT	AGGCGTTTTT
251	TGACCTCTTA	TCAAAAGGAG	CAATTAAAGG	TACTGTCTAA	TCCTGACCTG
	ACTGGAGAAT	AGTTTTCCCTC	GTTAATTTCC	ATGACAGATT	AGGACTGGAC
301	TTGGAATTTG	CTTCCGGTCT	GGTTCGCTTT	GAGGCTCGAA	TTGAAACGCG
	AACCTTAAAC	GAAGGCCAGA	CCAAGCGAAA	CTCCGAGCTT	AACTTTGCGC
351	ATATTTGAAG	TCTTTCGGGC	TTCCTCTTAA	TCTTTTTGAT	GCAATTTCGT
	TATAAACTTC	AGAAAGCCCG	AAGGAGAATT	AGAAAACTA	CGTTAAGCGA
401	TTGCTTCTGA	CTATAATAGA	CAGGGTAAAG	ACCTGATTTT	TGATTTATGG
	AACGAAGACT	GATATTATCT	GTCCCATTTC	TGGACTAAAA	ACTAAATACC
451	TCATTCTCGT	TTTCTGAACT	GTTTAAAGCA	TTTGAGGGGG	ATTCAATGAA
	AGTAAGAGCA	AAAGACTTGA	CAAATTTCTG	AAACTCCCC	TAAGTTACTT
501	TATTTATGAC	GATTCCGCAG	TATTGGACGC	TATCCAGTCT	AAACATTTTA
	ATAAATACTG	CTAAGGCGTC	ATAACCTGCG	ATAGGTCAGA	TTTGTAATAA
551	CAATTACCCC	CTCTGGCAAA	ACTTCCTTTG	CAAAAGCCTC	TCGCTATTTT
	GTTAATGGGG	GAGACCGTTT	TGAAGGAAAC	GTTTTTCGGAG	AGCGATAAAA
601	GGTTTCTATC	GTCGTCTGGT	TAATGAGGGT	TATGATAGTG	TTGCTCTTAC
	CCAAAGATAG	CAGCAGACCA	ATTACTCCCA	ATACTATCAC	AACGAGAATG
651	CATGCCTCGT	AATTCCTTTT	GGCGTTATGT	ATCTGCATTA	GTTGAGTGTG
	GTACGGAGCA	TTAAGGAAAA	CCGCAATACA	TAGACGTAAT	CAACTCACAC
701	GTATTCCTAA	ATCTCAATTG	ATGAATCTTT	CCACCTGTAA	TAATGTTGTT
	CATAAGGATT	TAGAGTTAAC	TACTTAGAAA	GGTGGACATT	ATTACAACAA
751	CCGTTAGTTC	GTTTTATTAA	CGTAGATTTT	TCCTCCCAAC	GTCCTGACTG
	GGCAATCAAG	CAAAATAATT	GCATCTAAAA	AGGAGGGTTG	CAGGACTGAC
801	GTATAATGAG	CCAGTTCTTA	AAATCGCATA	AGGTAATTCA	AAATGATTAA
	CATATTACTC	GGTCAAGAAT	TTTAGCGTAT	TCCATTAGT	TTTACTAATT
851	AGTTGAAATT	AAACCGTCTC	AAGCGCAATT	TACTACCCGT	TCTGGTGTTT
	TCAACTTTAA	TTTGGCAGAG	TTCGCGTTAA	ATGATGGGCA	AGACCACAAA

16/39

901	CTCGTCAGGG	CAAGCCTTAT	TCACTGAATG	AGCAGCTTTG	TTACGTTGAT
	GAGCAGTCCC	GTTCGGAATA	AGTGACTTAC	TCGTGCGAAAC	AATGCAACTA
951	TTGGGTAAATG	AATATCCGGT	GCTTGTCAAG	ATTACTCTCG	ACGAAGGTCA
	AACCCATTAC	TTATAGGCCA	CGAACAGTTC	TAATGAGAGC	TGCTTCCAGT
1001	GCCAGCGTAT	GCGCCTGGTC	TGTACACCGT	GCATCTGTCC	TCGTTCAAAG
	CGGTGCGATA	CGCGGACCAG	ACATGTGGCA	CGTAGACAGG	AGCAAGTTTC
1051	TTGGTCAGTT	CGGTTCTCTT	ATGATTGACC	GTCTGCGCCT	CGTTCCGGCT
	AACCAGTCAA	GCCAAGAGAA	TACTAACTGG	CAGACGCGGA	GCAAGGCCGA
1101	AAGTAACATG	GAGCAGGTCG	CGGATTTCTGA	CACAATTTAT	CAGGCGATGA
	TTTATTTGTAC	CTCGTCCAGC	GCCTAAAGCT	GTGTTAAATA	GTCCGCTACT
1151	TACAAATCTC	CGTTGTACTT	TGTTTCGCGC	TTGGTATAAT	CGCTGGGGGT
	ATGTTTAGAG	GCAACATGAA	ACAAAGCGCG	AACCATATTA	GCGACCCCCA
1201	CAAAGATGAG	TGTTTTAGTG	TATTCTTTTCG	CCTCTTTTCGT	TTTAGGTTGG
	GTTTCTACTC	ACAAAATCAC	ATAAGAAAGC	GGAGAAAGCA	AAATCCAACC
1251	TGCCTTCGTA	GTGGCATTAC	GTATTTTACC	CGTTTAATGG	AAACTTCCTC
	ACGGAAGCAT	CACCGTAATG	CATAAAATGG	GCAAATTACC	TTTGAAGGAG
1301	ATGCGTAAGT	CTTTAGTCCT	CAAAGCCTCC	GTAGCCGTTG	CTACCCTCGT
	TACGCATTCA	GAAATCAGGA	GT'TTCGGAGG	CATCGGCAAC	GATGGGAGCA
1351	TCCGATGCTG	TCTTTCGCTG	CTGAGGGTGA	CGATCCCGCA	AAAGCGGCCCT
	AGGCTACGAC	AGAAAGCGAC	GACTCCCACT	GCTAGGGCGT	TTTCGCCCGA
1401	TTGACTCCCT	GCAAGCCTCA	GCGACCGAAT	ATATCGGTTA	TGCGTGGGCG
	AACTGAGGGA	CGTTCGGAGT	CGCTGGCTTA	TATAGCCAAT	ACGCACCCGC
1451	ATGGTTGTTG	TCATTGTCGG	CGCAACTATC	GGTATCAAGC	TGTTTAAGAA
	TACCAACAAC	AGTAACAGCC	GCGTTGATAG	CCATAGTTCG	ACAAATTCTT
1501	ATTACCTCG	AAAGCAAGCT	GATAAAGGAG	GTTTCTCGAT	CGAGACGTTN
	TAAGTGAGGC	TTTCGTTCGA	CTATTTCCCTC	CAAAGAGCTA	GCTCTGCAAN
1551	NNNGAGGTTT	CAACTTTCAC	CATAATGAAA	TAAGATCACT	ACCGGGCGTA
	NNNCTCCAAG	GTTGAAAGTG	GTATTACTTT	ATTCTAGTGA	TGGCCCGCAT
1601	TTTTTTGAGT	TATCGAGATT	TTTCAAGGAGCT	AAGGAAGCTA	AAATGGAGAA
	AAAAAACTCA	ATAGCTCTAA	AAGTCCTCGA	TTCTTTCGAT	TTTACCTCTT
1651	AAAAATCACT	GGATATACCA	CCGTTGATAT	ATCCCAATGG	CATCGTAAAG
	TTTTTTAGTGA	CCTATATGGT	GGCAACTATA	TAGGGTTACC	GTAGCATTTT
1701	AACATTTTGA	GGCATTTTCAG	TCAGTTGCTC	AATGTACCTA	TAACCAGACC
	TTGTAAAAC	CCGTAAAGTC	AGTCAACGAG	TTACATGGAT	ATTGGTCTGG
1751	GTTCAGCTGG	ATATTACGGC	CTTTTTTAAAG	ACCGTAAAGA	AAAATAAGCA
	CAAGTCGACC	TATAATGCCG	GAAAAATTTT	TGGCATTTCT	TTTTATTTCGT

17/39

1801	CAAGTTTTAT	CCGGCCTTTA	TTCACATTCT	TGCCCCGCTG	ATGAATGCTC
	GTTCAAAATA	GGCCGGAAAT	AAGTGTAAGA	ACGGGCGGAC	TACTTACGAG
1851	ATCCGGAGTT	CCGTATGGCA	ATGAAAGACG	GTGAGCTGGT	GATATGGGAT
	TAGGCCTCAA	GGCATACCGT	TACTTTCTGC	CACTCGACCA	CTATACCCTA
1901	AGTGTTTACC	CTTGTTACAC	CGTTTTTCCAT	GAGCAAACCTG	AAACGTTTTTC
	TCACAAGTGG	GAACAATGTG	GCAAAAGGTA	CTCGTTTGAC	TTTGCAAAAG
1951	ATCGCTCTGG	AGTGAATACC	ACGACGATTT	CCGGCAGTTT	CTACACATAT
	TAGCGAGACC	TCACTTATGG	TGCTGCTAAA	GGCCGTCAAA	GATGTGTATA
2001	ATTCGCAAGA	TGTGGCGTGT	TACGGTGAAA	ACCTGGCCTA	TTTCCCTAAA
	TAAGCGTTCT	ACACCGCACA	ATGCCACTTT	TGGACCGGAT	AAAGGGATTT
2051	GGGTTTATTG	AGAATATGTT	TTTCGTCTCA	GCCAATCCCT	GGGTGAGTTT
	CCCAAATAAC	TCTTATACAA	AAAGCAGAGT	CGGTTAGGGA	CCCCTCAAA
2101	CACCAAGTTTT	GATTTAAACG	TAGCCAATAT	GGACAACCTC	TTCGCCCCCG
	GTGGTCAAAA	CTAAATTTGC	ATCGGTTATA	CCTGTTGAAG	AAGCGGGGGC
2151	TTTTCACTAT	GGGCAAATAT	TATACGCAAG	GCGACAAGGT	GCTGATGCCG
	AAAAGTGATA	CCCGTTTATA	ATATGCGTTC	CGCTGTTCCA	CGACTACGGC
2201	CTGGCGATTG	AGGTTTCATCA	TGCCGTTTGT	GATGGCTTCC	ATGTCGGCAG
	GACCGCTAAG	TCCAAGTAGT	ACGGCAAACA	CTACCGAAGG	TACAGCCGTC
2251	AATGCTTAAT	GAATTACAAC	AGTACTGCGA	TGAGTGGCAG	GGCGGGGCGT
	TTACGAATTA	CTTAATGTTG	TCATGACGCT	ACTCACCGTC	CCGCCCCGCA
2301	AATTTTTTTA	AGGCAGTTAT	TGGTGCCCTT	AAACGCCTGG	TGCTAGCCTG
	TTAAAAAAT	TCCGTCAATA	ACCACGGGAA	TTTGCGGACC	ACGATCGGAC
2351	AGGCCAGTTT	GCTCAGGCTC	TCCCCGTGGA	GGTAATAATT	GCTCGACCGA
	TCCGGTCAAA	CGAGTCCGAG	AGGGGCACCT	CCATTATTAA	CGAGCTGGCT
2401	TAAAAGCGGC	TTCCTGACAG	GAGGCCGTTT	TGTTTTGCAG	CCCACCTCAA
	ATTTTCGCCG	AAGGACTGTC	CTCCGGCAAA	ACAAAACGTC	GGGTGGAGTT
2451	CGCAATTAAT	GTGAGTTAGC	TCACTCATTA	GGCACCCCAG	GCTTTTACACT
	GCGTTAATTA	CACTCAATCG	AGTGAGTAAT	CCGTGGGGTC	CGAAATGTGA
2501	TTATGCTTCC	GGCTCGTATG	TTGTGTGGAA	TTGTGAGCGG	ATAACAATTT
	AATACGAAGG	CCGAGCATAC	AACACACCTT	AACACTCGCC	TATTGTTAAA
2551	CACACAGGAA	ACAGCTATGA	CCATGATTAC	GAATTTCTAG	ATAACGAGGG
	GTGTGTCCTT	TGTCGATACT	GGTACTAATG	CTTAAAGATC	TATTGCTCCC
2601	CAAAAAATGA	AAAAGACAGC	TATCGCGATT	GCAGTGGCAC	TGGCTGGTTT
	GTTTTTTACT	TTTTCTGTCT	ATAGCGCTAA	CGTCACCGTG	ACCGACCAAA
2651	CGCTACCGTA	GCGCAGGCCG	ACTACAAAGA	TGTCGACGCC	GGTGGTCCGA
	GCGATGGCAT	CGCGTCCGGC	TGATGTTTCT	ACAGCTGCGG	CCACCAGCCT

18/39

2701 TCGCCCGGCT AGAGGAAAAA GTGAAAACCT TGAAAGCGCA AAACCTCCGAG
AGCGGGCCGA TCTCCTTTTT CACTTTTGGA ACTTTCGCGT TTTGAGGCTC

2751 CTGGCGTCCA CGGCCAACAT GCTCAGGGAA CAGGTGGCAC AGCTTAAACA
GACCGCAGGT GCCGGTTGTA CGAGTCCCTT GTCCACCGTG TCGAATTTGT

EcoRI

~~~~~

2801 GAAAGTCATG AACCACGGTG GTGCCGAATT CAATGCTGGC GGCGGCTCTG  
CTTTCAGTAC TTGGTGCCAC CACGGCTTAA GTTACGACCG CCGCCGAGAC

2851 GTGGTGGTTC TGGTGGCGGC TCTGAGGGTG GTGGCTCTGA GGGTGGCGGT  
CACCACCAAG ACCACCGCCG AGACTCCCAC CACCGAGACT CCCACCGCCA

2901 TCTGAGGGTG GCGGCTCTGA GGGAGGCGGT TCCGGTGGTG GCTCTGGTTC  
AGACTCCCAC CGCCGAGACT CCCTCCGCCA AGGCCACCAC CGAGACCAAG

2951 CGGTGATTTT GATTATGAAA AGATGGCAAA CGCTAATAAG GGGGCTATGA  
GCCACTAAAA CTAATACTTT TCTACCGTTT GCGATTATTC CCCCATACT

3001 CCGAAAATGC CGATGAAAAC GCGCTACAGT CTGACGCTAA AGGCAAACCTT  
GGCTTTTACG GCTACTTTTG CGCGATGTCA GACTGCGATT TCCGTTTGAA

Clal

~~~~~

3051 GATTCTGTCTG CTAATGATTA CGGTGCTGCT ATCGATGGTT TCATTGGTGA
CTAAGACAGC GATGACTAAT GCCACGACGA TAGCTACCAA AGTAACCACT

3101 CGTTTCCGGC CTTGCTAATG GTAATGGTGC TACTGGTGAT TTTGCTGGCT
GCAAAGGCCG GAACGATTAC CATTACCACG ATGACCACTA AAACGACCGA

3151 CTAATTCCCA AATGGCTCAA GTCGGTGACG GTGATAATTC ACCTTTAATG
GATTAAGGGT TTACCGAGTT CAGCCACTGC CACTATTAAG TGGAAATTAC

3201 AATAATTTCC GTCAATATTT ACCTTCCCTC CCTCAATCGG TTGAATGTCTG
TTATTAAAGG CAGTTATAAA TGGAAGGGAG GGAGTTAGCC AACTTACAGC

3251 CCCTTTTGTC TTTAGCGCTG GTAAACCATA TGAATTTTCT ATTGATTGTG
GGGAAAACAG AAATCGCGAC CATTTGGTAT ACTTAAAAGA TAACTAACAC

3301 ACAAATAAAA CTTATTCCGT GGTGTCTTTG CGTTTCTTTT ATATGTTGCC
TGTTTTATTT GAATAAGGCA CCACAGAAAC GCAAAGAAAA TATACAACGG

3351 ACCTTTATGT ATGTATTTTC TACGTTTGCT AACATACTGC GTAATAAGGA
TGGAAATACA TACATAAAAG ATGCAAACGA TTGTATGACG CATTATTCCT

HindIII

~~~~~

3401 GTCTTGATAA GCTTCGAGAA ATTCACCTCG AAAGCAAGCT GATAAACCGA  
CAGAACTATT CGAAGCTCTT TAAGTGGAGC TTTCGTTCTGA CTATTTGGCT

3451 TACAATTAAA GGCTCCTTTT GGAGCCTTTT TTTTGGGAGA ATTAATTCAA  
ATGTTAATTT CCGAGGAAAA CCTCGGAAAA AAAAACCTCT TAATTAAGTT

3501 TCATGCCAGT TCTTTTGGGT ATTCGTTTAT TATTGCGTTT CCTCGGTTTC

19/39

|      |                          |                          |                          |                           |                           |
|------|--------------------------|--------------------------|--------------------------|---------------------------|---------------------------|
|      | AGTACGGTCA               | AGAAAACCCA               | TAAGGCAATA               | ATAACGCAAA                | GGAGCCAAAG                |
| 3551 | CTTCTGGTAA<br>GAAGACCATT | CTTTGTTCGG<br>GAAACAAGCC | CTATCTGCTT<br>GATAGACGAA | ACTTTCCTTA<br>TGAAAGGAAT  | AAAAGGGCTT<br>TTTTCCCGAA  |
| 3601 | CGGTAAGATA<br>GCCATTCTAT | GCTATTGCTA<br>CGATAACGAT | TTTCATTGTT<br>AAAGTAACAA | TCTTGCTCTT<br>AGAACGAGAA  | ATTATTGGGC<br>TAATAACCCG  |
| 3651 | TTAACTCAAT<br>AATTGAGTTA | TCTTGTGGGT<br>AGAACACCCA | TATCTCTCTG<br>ATAGAGAGAC | ATATTAGCGC<br>TATAATCGCG  | ACAAATTACCC<br>TGTTAATGGG |
| 3701 | TCTGATTTTG<br>AGACTAAAAC | TTCAGGGCGT<br>AAGTCCCGCA | TCAGTTAATT<br>AGTCAATTAA | CTCCCGTCTA<br>GAGGGCAGAT  | ATGCGCTTCC<br>TACGCGAAGG  |
| 3751 | CTGTTTTTAT<br>GACAAAAATA | GTTATTCTCT<br>CAATAAGAGA | CTGTAAAGGC<br>GACATTTCCG | TGCTATTTTC<br>ACGATAAAAAG | ATTTTTGACG<br>TAAAAACTGC  |
| 3801 | TTAAACAAAA<br>AATTTGTTTT | AATCGTTTCT<br>TTAGCAAAGA | TATTTGGATT<br>ATAAACCTAA | GGGATAAATA<br>CCCTATTTAT  | AATATGGCTG<br>TTATACCGAC  |
| 3851 | TTTATTTTGT<br>AAATAAAACA | AACTGGCAAA<br>TTGACCGTTT | TTAGGCTCTG<br>AATCCGAGAC | GAAAGACGCT<br>CTTTCCTGCGA | CGTTAGCGTT<br>GCAATCGCAA  |
| 3901 | GGTAAGATTC<br>CCATTCTAAG | AGGATAAAAT<br>TCCTATTTTA | TGTAGCTGGG<br>ACATCGACCC | TGCAAAATAG<br>ACGTTTTATC  | CAACTAATCT<br>GTTGATTAGA  |
| 3951 | TGATTTAAGG<br>ACTAAATTCC | CTTCAAAACC<br>GAAGTTTGGG | TCCCGCAAGT<br>AGGGCGTTCA | CGGGAGGTTC<br>GCCCTCCAAG  | GCTAAAACGC<br>CGATTTTGCG  |
| 4001 | CTCGCGTTCT<br>GAGCGCAAGA | TAGAATACCG<br>ATCTTATGGC | GATAAGCCTT<br>CTATTCGGAA | CTATTTCTGA<br>GATAAAGACT  | TTTGCTTGCT<br>AAACGAACGA  |
| 4051 | ATTGGTCGTG<br>TAACCAGCAC | GTAATGATTC<br>CATTACTAAG | CTACGACGAA<br>GATGCTGCTT | AATAAAAACG<br>TTATTTTTGC  | GTTTGCTTGT<br>CAAACGAACA  |
| 4101 | TCTTGATGAA<br>AGAACTACTT | TGCGGTACTT<br>ACGCCATGAA | GGTTTAATAC<br>CCAAATTATG | CCGTTTCATG<br>GGCAAGTACC  | AATGACAAGG<br>TTACTGTTCC  |
| 4151 | AAAGACAGCC<br>TTTCTGTTCG | GATTATTGAT<br>CTAATAACTA | TGGTTTCTTC<br>ACCAAAGAAG | ATGCTCGTAA<br>TACGAGCATT  | ATTGGGATGG<br>TAACCC'TACC |
| 4201 | GATATTATTT<br>CTATAATAAA | TTCTTGTTCA<br>AAGAACAAGT | GGATTTATCT<br>CCTAAATAGA | ATTGTTGATA<br>TAACAACATAT | AACAGGCGCG<br>TTGTCCGCGC  |
| 4251 | TTCTGCATTA<br>AAGACGTAAT | GCTGAACACG<br>CGACTTGTGC | TTGTTTATTG<br>AACAAATAAC | TCGCCGTCTG<br>AGCGGCAGAC  | GACAGAATTA<br>CTGTCTTAAT  |
| 4301 | CTTTACCCTT<br>GAAATGGGAA | TGTCGGCACT<br>ACAGCCGTGA | TTATATTCTC<br>AATATAAGAG | TTGTTACTGG<br>AACAAATGACC | CTCAAAAATG<br>GAGTTTTTAC  |
| 4351 | CCTCTGCCTA<br>GGAGACGGAT | AATTACATGT<br>TTAATGTACA | TGGTGTGTTT<br>ACCACAACAA | AAATATGGTG<br>TTTATACCAC  | ATTCTCAATT<br>TAAGAGTTAA  |

20/39

|      |                           |                           |                            |                           |                           |
|------|---------------------------|---------------------------|----------------------------|---------------------------|---------------------------|
| 4401 | AAGCCCTACT<br>TTCGGGATGA  | GTTGAGCGTT<br>CAACTCGCAA  | GGCTTTATAC<br>CCGAAATATG   | TGGTAAGAAT<br>ACCATTCCTTA | TTATATAACG<br>AATATATTGC  |
| 4451 | CATATGACAC<br>GTATACTGTG  | TAAACAGGCT<br>ATTTGTCCGA  | TTTTCCAGTA<br>AAAAGGTCAT   | ATTATGATTTC<br>TAATACTAAG | AGGTGTTTAT<br>TCCACAAATA  |
| 4501 | TCATATTTAA<br>AGTATAAATT  | CCCCTTATTT<br>GGGGAATAAA  | ATCACACGGT<br>TAGTGTGCCA   | CGGTATTTCA<br>GCCATAAAGT  | AACCATTAAA<br>TTGGTAATTT  |
| 4551 | TTTAGGTCAG<br>AAATCCAGTC  | AAGATGAAAT<br>TTCTACTTTA  | TAACTAAAAAT<br>ATTGATTTTA  | ATATTTGAAA<br>TATAAACTTT  | AAGTTTTCTC<br>TTCAAAAGAG  |
| 4601 | GCGTTCCTTG<br>CGCAAGAAAC  | TCTTGCGATA<br>AGAACGCTAT  | GGATTTGCAT<br>CCTAAACGTA   | CAGCATTTAC<br>GTCGTAAATG  | ATATAGTTAT<br>TATATCAATA  |
| 4651 | ATAACCCAAC<br>TATTGGGTTG  | CTAAGCCGGA<br>GATTCGGCCT  | GGTTAAAAAG<br>CCAATTTTTC   | GTAGTCTCTC<br>CATCAGAGAG  | AGACCTATGA<br>TCTGGATACT  |
| 4701 | TTTTGATAAA<br>AAAAC TATTT | TTCACTATTG<br>AAGTGATAAC  | ACTCTTCTCA<br>TGAGAAGAGT   | GCGTCTTAAT<br>CGCAGAATTA  | CTAAGCTATC<br>GATTCGATAG  |
| 4751 | GCTATGTTTT<br>CGATACAAA   | CAAGGATTCT<br>GTTCCCTAAGA | AAGGGAAAAT<br>TTCCCTTTTA   | TAATTAATAG<br>ATTAATTATC  | CGACGATTTA<br>GCTGCTAAAT  |
| 4801 | CAGAAGCAAG<br>GTCTTCGTTT  | GTTATTCCAT<br>CAATAAGGTA  | CACATATATT<br>GTGTATATAA   | GATTTATGTA<br>CTAAATACAT  | CTGTTTCAAT<br>GACAAAAGTTA |
| 4851 | TAAAAAAGGT<br>ATTTTTTTCCA | AATTCAAATG<br>TTAAGTTTAC  | AAATTGTTAA<br>TTTAACAATT   | ATGTAATTAA<br>TACATTAATT  | TTTTGTTTTTC<br>AAAACAAAAG |
| 4901 | TTGATGTTTG<br>AACTACAAAC  | TTTCATCATC<br>AAAGTAGTAG  | TTCTTTTGCT<br>AAGAAAACGA   | CAAGTAATTG<br>GTTTCATTAAC | AAATGAATAA<br>TTTACTTATT  |
| 4951 | TTCGCCTCTG<br>AAGCGGAGAC  | CGCGATTTTCG<br>GCGCTAAAGC | TGACTTG GTA<br>ACTGAACCAT  | TTCAAAGCAA<br>AAGTTTCGTT  | ACAGGTGAAT<br>TGTCCACTTA  |
| 5001 | CTGTTATTGT<br>GACAATAACA  | CTCACCTGAT<br>GAGTGGACTA  | GTAAAGGTA<br>CAATTTCCAT    | CAGTGACTGT<br>GTCAC TGACA | ATATTCCTCT<br>TATAAGGAGA  |
| 5051 | GACGTTAAGC<br>CTGCAATTCG  | CTGAAAATTT<br>GACTTTTAA   | ACGCAATTTTC<br>TGC GTTAAAG | TTTATCTCTG<br>AAATAGAGAC  | TTTTACGTGC<br>AAAATGCACG  |
| 5101 | TAATAATTTT<br>ATTATTAAAA  | GATATGGTTG<br>CTATACCAAC  | GCTCAATTCC<br>CGAGTTAAGG   | TTCCATAATT<br>AAGGTATTAA  | CAGAAATATA<br>GTCTTTTATAT |
| 5151 | ACCCAAATAG<br>TGGGTTTATC  | TCAGGATTAT<br>AGTCCTAATA  | ATTGATGAAT<br>TAATACTTA    | TGCCATCATC<br>ACGGTAGTAG  | TGATATTTCAG<br>ACTATAAGTC |
| 5201 | GAATATGATG<br>CTTATACTAC  | ATAATTCCGC<br>TATTAAGGCG  | TCCTTCTGGT<br>AGGAAGACCA   | GGTTTCTTTG<br>CCAAAGAAAC  | TTCCGCAAAA<br>AAGGCGTTTT  |
| 5251 | TGATAATGTT<br>ACTATTACAA  | ACTCAAACAT<br>TGAGTTTGTA  | TTAAAATTAA<br>AATTTTAATT   | TAACGTTTCG<br>ATTGCAAGCG  | GCAAAGGATT<br>CGTTTCCTAA  |

21/39

|      |            |            |            |            |            |
|------|------------|------------|------------|------------|------------|
| 5301 | TAATAAGGGT | TGTAGAATTG | TTTGTAAAT  | CTAATACATC | TAAATCCTCA |
|      | ATTATTCCCA | ACATCTTAAC | AAACAATTTA | GATTATGTAG | ATTTAGGAGT |
| 5351 | AATGTATTAT | CTGTTGATGG | TTCTAACTTA | TTAGTAGTTA | GCGCCCCTAA |
|      | TTACATAATA | GACAACTACC | AAGATTGAAT | AATCATCAAT | CGCGGGGATT |
| 5401 | AGATATTTTA | GATAACCTTC | CGCAATTTCT | TTCTACTGTT | GATTTGCCAA |
|      | TCTATAAAAT | CTATTGGAAG | GCGTTAAAGA | AAGATGACAA | CTAAACGGTT |
| 5451 | CTGACCAGAT | ATTGATTGAA | GGATTAATTT | TCGAGGTTCA | GCAAGGTGAT |
|      | GACTGGTCTA | TAACTAACTT | CCTAATTAAA | AGCTCCAAGT | CGTTCCACTA |
| 5501 | GCTTTAGATT | TTTCCTTTGC | TGCTGGCTCT | CAGCGCGGCA | CTGTTGCTGG |
|      | CGAAATCTAA | AAAGGAAACG | ACGACCGAGA | GTCGCGCCGT | GACAACGACC |
| 5551 | TGGTGTAAAT | ACTGACCGTC | TAACCTCTGT | TTTATCTTCT | GCGGGTGGTT |
|      | ACCACAATTA | TGACTGGCAG | ATTGGAGACA | AAATAGAAGA | CGCCCACCAA |
| 5601 | CGTTCGGTAT | TTTTAACGGC | GATGTTTTAG | GGCTATCAGT | TCGCGCATTA |
|      | GCAAGCCATA | AAAATTGCCG | CTACAAAATC | CCGATAGTCA | AGCGCGTAAT |
| 5651 | AAGACTAATA | GCCATTCAAA | AATATTGTCT | GTGCCTCGTA | TTCTTACGCT |
|      | TTCTGATTAT | CGGTAAGTTT | TTATAACAGA | CACGGAGCAT | AAGAATGCGA |
| 5701 | TTCAGGTCAG | AAGGGTTCTA | TTTCTGTTGG | CCAGAATGTC | CCTTTTATTA |
|      | AAGTCCAGTC | TTCCCAAGAT | AAAGACAACC | GGTCTTACAG | GGAAAATAAT |
| 5751 | CTGGTCGTGT | AACTGGTGAA | TCTGCCAATG | TAAATAATCC | ATTTCAGACG |
|      | GACCAGCACA | TTGACCACTT | AGACGGTTAC | ATTTATTAGG | TAAAGTCTGC |
| 5801 | GTTGAGCGTC | AAAATGTTGG | TATTTCTATG | AGTGTTTTTC | CCGTTGCAAT |
|      | CAACTCGCAG | TTTTACAACC | ATAAAGATAC | TCACAAAAG  | GGCAACGTTA |
| 5851 | GGCTGGCGGT | AATATTGTTT | TAGATATAAC | CAGTAAGGCC | GATAGTTTGA |
|      | CCGACCGCCA | TTATAACAAA | ATCTATATTG | GTCATTCCGG | CTATCAAAC  |
| 5901 | GTTCTTCTAC | TCAGGCAAGT | GATGTTATTA | CTAATCAAAG | AAGTATTGCG |
|      | CAAGAAGATG | AGTCCGTTCA | CTACAATAAT | GATTAGTTTC | TTCATAACGC |
| 5951 | ACAACGGTTA | ATTTGCGTGA | TGGTCAGACT | CTTTTGCTCG | GTGGCCTCAC |
|      | TGTTGCCAAT | TAAACGCACT | ACCAGTCTGA | GAAAACGAGC | CACCGGAGTG |
| 6001 | TGATTACAAA | AACACTTCTC | AAGATTCTGG | TGTGCCGTTT | CTGTCTAAAA |
|      | ACTAATGTTT | TTGTGAAGAG | TTCTAAGACC | ACACGGCAAG | GACAGATTTT |
| 6051 | TCCCTTTAAT | CGGCCTCCTG | TTTAGCTCCC | GTTCTGATTC | TAACGAGGAA |
|      | AGGGAAATTA | GCCGGAGGAC | AAATCGAGGG | CAAGACTAAG | ATTGCTCCTT |
| 6101 | AGCACGTTGT | ACGTGCTCGT | CAAAGCAACC | ATAGTACGCG | CCCTGTAGCG |
|      | TCGTGCAACA | TGCACGAGCA | GTTTCGTTGG | TATCATGCGC | GGGACATCGC |
| 6151 | GCGCATTAAG | CGCGGCGGGT | GTGGTGGTTA | CGCGCAGCGT | GACCGCTACA |
|      | CGCGTAATTC | GCGCCGCCCA | CACCACCAAT | GCGCGTCGCA | CTGGCGATGT |

22/39

6201 CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC CTTCTTTTCT  
GAACGGTCGC GGGATCGCGG GCGAGGAAAG CGAAAGAAGG GAAGGAAAGA

BamHI

~~~~~

6251 CGCCACGTTC TCCGGCTTTC CCCGTCAAGC TCTAAATCGG GGGATCCCTT
GCGGTGCAAG AGGCCGAAAG GGGCAGTTTC AGATTTAGCC CCTAGGGAA

6301 TAGGGTTCCG ATTTAGTGCT TTACGGCACC TCGACCTCCA AAAACTTGAT
ATCCCAAGGC TAAATCACGA AATGCCGTGG AGCTGGAGGT TTTTGAAC TA

6351 TTGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTTCG
AACCCACTAC CAAGTGCATC ACCCGGTAGC GGGACTATCT GCCAAAAAGC

6401 CCCTTTGACG TTGGAGTCCA CGTTCCTTAA TAGTGGACTC TTGTTCCAAA
GGGAAACTGC AACCTCAGGT GCAAGAAATT ATCACCTGAG AACAAGGTTT

6451 CTGGAACAAC ACTCACAAC AACTCGGCCT ATTCTTTTGA TTTATAAGGA
GACCTTGTTG TGAGTGTTGA TTGAGCCGGA TAAGAAACT AAATATTCCT

6501 TTTTTGTCAT TTTCTGCTTA CTGGTTAAAA AATAAGCTGA TTTAACAAAT
AAAAACAGTA AAAGACGAAT GACCAATTTT TTATTTCGACT AAATTGTTTA

6551 ATTTAACGCG AAATTTAACA AAACATTAAC GTTTACAATT TAAATATTTG
TAAATTGCGC TTTAAATTGT TTTGTAATTG CAAATGTAA ATTTATAAAC

6601 CTTATACAAT CATCCTGTTT TTGGGGCTTT TCTGATTATC AACCGGGGTA
GAATATGTTA GTAGGACAAA AACCCCGAAA AGACTAATAG TTGGCCCCAT

ClaI

~~~~~

6651 CATATGATTG ACATGCTAGT TTTACGATTA CCGTTCATCG ATTCTCTTGT  
GTATACTAAC TGTACGATCA AAATGCTAAT GGCAAGTAGC TAAGAGAACA

6701 TTGCTCCAGA CTTTCAGGTA ATGACCTGAT AGCCTTTGTA GACCTCTCAA  
AACGAGGTCT GAAAGTCCAT TACTGGACTA TCGGAAACAT CTGGAGAGTT

6751 AAATAGCTAC CCTCTCCGGC ATGAATTTAT CAGCTAGAAC GGTTGAATAT  
TTTATCGATG GGAGAGGCCG TACTTAAATA GTCGATCTTG CCAACTTATA

6801 CATATTGACG GTGATTTGAC TGTCTCCGGC CTTTCTCACC CGTTTGAATC  
GTATAACTGC CACTAAACTG ACAGAGGCCG GAAAGAGTGG GCAAACCTAG

6851 TTTGCCTACT CATTACTCCG GCATTGCATT TAAAATATAT GAGGGTTCTA  
AAACGGATGA GTAATGAGGC CGTAACGTAA ATTTTATATA CTCCCAAGAT

6901 AAAATTTTTA TCCCTGCGTT GAAATTAAGG CTTCAACGAG AAAAGTATTA  
TTTTAAAAAT AGGGACGCAA CTTTAATTCC GAAGTGGTCG TTTTCATAAT

6951 CAGGGTCATA ATGTTTTTGG TACAACCGAT TTAGCTTTAT GCTCTGAGGC  
GTCCCAGTAT TACAAAAACC ATGTTGGCTA AATCGAAATA CGAGACTCCG

7001 TTTATTGCTT AATTTTGCTA ACTCTCTGCC TTGCTTGTA GATTATTTGG

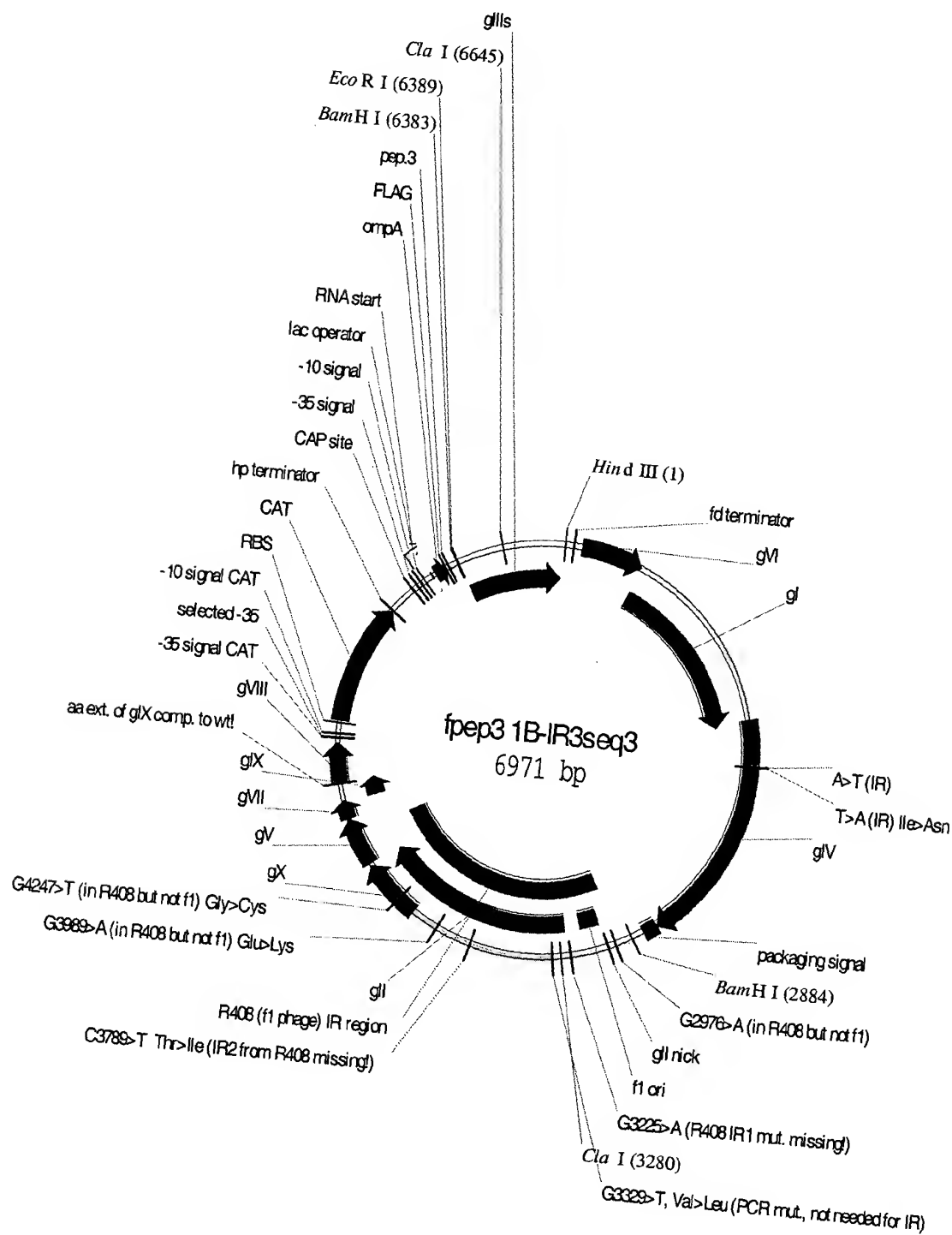
23/39

AAATAACGAA TTAAAACGAT TGAGAGACGG AACGAACATG CTAAATAACC

7051 ATGTT  
TACAA

**Figure 4**

24/39





25/39

HindIII

~~~~~

1	AGCTTCGAGA	AATTCACCTC	GAAAGCAAGC	TGATAAACCG	ATACAATTAA
	TCGAAGCTCT	TTAAGTGGAG	CTTTCGTTTC	ACTATTTGGC	TATGTTAATT
51	AGGCTCCTTT	TGGAGCCTTT	TTTTTTGGAG	AATTAATTCA	ATCATGCCAG
	TCCGAGGAAA	ACCTCGGAAA	AAAAAACCTC	TTAATTAAGT	TAGTACGGTC
101	TTCTTTTGGG	TATTCGGTTA	TTATTGCGTT	TCCTCGGTTT	CCTTCTGGTA
	AAGAAAACCC	ATAAGGCAAT	AATAACGCAA	AGGAGCCAAA	GGAAGACCAT
151	ACTTTGTTTC	GCTATCTGCT	TACTTTCCTT	AAAAAGGGCT	TCGGTAAGAT
	TGAAACAAGC	CGATAGACGA	ATGAAAGGAA	TTTTTCCCGA	AGCCATTCTA
201	AGCTATTGCT	ATTTTCATTGT	TTCTTGCTCT	TATTATTGGG	CTTAACCTCA
	TCGATAACGA	TAAAGTAACA	AAGAACGAGA	ATAATAACCC	GAATTGAGTT
251	TTCTTGTTGG	TTATCTCTCT	GATATTAGCG	CACAATTACC	CTCTGATTTT
	AAGAACACCC	AATAGAGAGA	CTATAATCGC	GTGTTAATGG	GAGACTAAAA
301	G TTCAGGGCG	TTCAGTTAAT	TCTCCCGTCT	AATGCGCTTC	CCTGTTTTTT
	CAAGTCCCGC	AAGTCAATTA	AGAGGGCAGA	TTACGCGAAG	GGACAAAAAT
351	TGTTATTCTC	TCTGTAAAGG	CTGCTATTTT	CATTTTTGAC	GTTAAACAAA
	ACAATAAGAG	AGACATTTCC	GACGATAAAA	GTAAAAACTG	CAATTTGTTT
401	AAATCGTTTC	TTATTTGGAT	TGGGATAAAT	AAATATGGCT	GTTTATTTTG
	TTTAGCAAAG	AATAAACCTA	ACCCTATTTA	TTTATACCGA	CAAATAAAAC
451	TAAGTGGCAA	ATTAGGCTCT	GGAAAGACGC	TCGTTAGCGT	TGGTAAGATT
	ATTGACCGTT	TAATCCGAGA	CCTTTCTGCG	AGCAATCGCA	ACCATTCTAA
501	CAGGATAAAA	TTGTAGCTGG	GTGCAAAATA	GCAACTAATC	TTGATTTAAG
	GTCCTATTTT	AACATCGACC	CACGTTTTAT	CGTTGATTAG	AACTAAATTC
551	GCTTCAAAAC	CTCCCGCAAG	TCGGGAGGTT	CGCTAAAACG	CCTCGCGTTC
	CGAAGTTTTG	GAGGGCGTTC	AGCCCTCCAA	GCGATTTTGC	GGAGCGCAAG
601	TTAGAATACC	GGATAAGCCT	TCTATTTCTG	ATTTGCTTGC	TATTGGTCGT
	AATCTTATGG	CCTATTCGGA	AGATAAAGAC	TAAACGAACG	ATAACCAGCA
651	GGTAATGATT	CCTACGACGA	AAATAAAAAC	GGTTTGCTTG	TTCTTGATGA
	CCATTACTAA	GGATGCTGCT	TTTATTTTTG	CCAAACGAAC	AAGAACTACT
701	ATGCGGTACT	TGGTTTAATA	CCCGTTCATG	GAATGACAAG	GAAAGACAGC
	TACGCCATGA	ACCAAATTAT	GGGCAAGTAC	CTTACTGTTC	CTTCTGTGTC
751	CGATTATTGA	TTGGTTTCTT	CATGCTCGTA	AATTGGGATG	GGATATTATT
	GCTAATAACT	AACCAAAGAA	GTACGAGCAT	TTAACCTTAC	CCTATAATAA
801	TTTCTTGTTT	AGGATTTATC	TATTGTTGAT	AAACAGGCGC	GTTCTGCATT
	AAAGAACAAG	TCCTAAATAG	ATAACAATA	TTTGTCCGCG	CAAGACGTAA

26/39

851	AGCTGAACAC TCGACTTGTG	GTTGTTTATT CAACAAATAA	GTCGCCGTCT CAGCGGCAGA	GGACAGAATT CCTGTCTTAA	ACTTTACCCT TGAAATGGGA
901	TTGTCCGGCAC AACAGCCGTG	TTTATATTCT AAATATAAGA	CTTGTTACTG GAACAATGAC	GCTCAAAAAT CGAGTTTTTA	GCCTCTGCCT CGGAGACGGA
951	AAATTACATG TTTAATGTAC	TTGGTGTTGT AACCACAACA	TAAATATGGT ATTTATACCA	GATTCTCAAT CTAAGAGTTA	TAAGCCCTAC ATTCGGGATG
1001	TGTTGAGCGT ACAACCTCGCA	TGGCTTTATA ACCGAAATAT	CTGGTAAGAA GACCATCTT	TTTATATAAC AAATATATTG	GCATATGACA CGTATACTGT
1051	CTAAACAGGC GATTTGTCCG	TTTTTCCAGT AAAAAGGTCA	AATTATGATT TTAATACTAA	CAGGTGTTTA GTCCACAAAT	TTCATATTTA AAGTATAAAT
1101	ACCCCTTATT TGGGGAATAA	TATCACACGG ATAGTGTGCC	TCGGTATTTT AGCCATAAAG	AAACCATTAA TTTGGTAATT	ATTTAGGTCA TAAATCCAGT
1151	GAAGATGAAA CTTCTACTTT	TTAACTAAAA AATTGATTTT	TATATTTGAA ATATAAACTT	AAAGTTTTCT TTTCAAAAGA	CGCGTTCTTT GCGCAAGAAA
1201	GTCTTGCGAT CAGAACGCTA	AGGATTTGCA TCCTAAACGT	TCAGCATTTA AGTCGTAAAT	CATATAGTTA GTATATCAAT	TATAACCCAA ATATTGGGTT
1251	CCTAAGCCGG GGATTCGGCC	AGGTTAAAAA TCCAATTTTT	GGTAGTCTCT CCATCAGAGA	CAGACCTATG GTCTGGATAC	ATTTTGATAA TAAACTATT
1301	ATTCACTATT TAAGTGATAA	GACTCTTCTC CTGAGAAGAG	AGCGTCTTAA TCGCAGAATT	TCTAAGCTAT AGATTTCGATA	CGCTATGTTT GCGATACAAA
1351	TCAAGGATTC AGTTCCTAAG	TAAGGGAAAA ATTCCCTTTT	TTAATTAATA AATTAATTAT	GCGACGATTT CGCTGCTAAA	ACAGAAGCAA TGTCTTCGTT
1401	GGTTATTCCA CCAATAAGGT	TCACATATAT AGTGTATATA	TGATTTATGT ACTAAATACA	ACTGTTTCAA TGACAAAGTT	TTAAAAAAGG AATTTTTTCC
1451	TAATTCAAAT ATTAAGTTTA	GAAATTGTTA CTTTAACAAT	AATGTAATTA TTACATTAAT	ATTTTGTTTT TAAACAAAA	CTTGATGTTT GACTACAAA
1501	GTTTCATCAT CAAAGTAGTA	CTTCTTTTGC GAAGAAAACG	TCAAGTAATT AGTTCATTAA	GAAATGAATA CTTTACTTAT	ATTCGCCTCT TAAGCGGAGA
1551	GCGCGATTTT CGCGCTAAAG	GTGACTTGGT CACTGAACCA	ATTCAAAGCA TAAGTTTCGT	AACAGGTGAA TTGTCCACTT	TCTGTTATTG AGACAATAAC
1601	TCTCACCTGA AGAGTGGACT	TGTTAAAGGT ACAATTTCCA	ACAGTGACTG TGTCACTGAC	TATATTCCTC ATATAAGGAG	TGACGTTAAG ACTGCAATTC
1651	CCTGAAAATT GGACTTTTAA	TACGCAATTT ATGCGTTAAA	CTTTATCTCT GAAATAGAGA	GTTTTACGTG CAAAATGCAC	CTAATAATTT GATTATTAAA
1701	TGATATGGTT ACTATACCAA	GGCTCTAATC CCGAGATTAG	CTTCCATAAT GAAGGTATTA	TCAGAAATAT AGTCTTTATA	AACCCAAATA TTGGGTTTAT

27/39

1751	GTCAGGATTA	TATTGATGAA	TTGCCATCAT	CTGATATTCA	GGAATATGAT
	CAGTCCTAAT	ATAACTACTT	AACGGTAGTA	GACTATAAGT	CCTTATACTA
1801	GATAATTCCG	CTCCTTCTGG	TGGTTTCTTT	GTTCCGCAAA	ATGATAATGT
	CTATTAAGGC	GAGGAAGACC	ACCAAAGAAA	CAAGGCGTTT	TACTATTACA
1851	TACTCAAACA	TTTAAAATTA	ATAACGTTCG	CGCAAAGGAT	TTAATAAGGG
	ATGAGTTTGT	AAATTTTAAT	TATTGCAAGC	GCGTTTCCTA	AATTATTCCC
1901	TTGTAGAATT	GTTTGTAA	TCTAATACAT	CTAAATCCTC	AAATGTATTA
	AACATCTTAA	CAAACAATTT	AGATTATGTA	GATTTAGGAG	TTTACATAAT
1951	TCTGTTGATG	GTTCTAACTT	ATTAGTAGTT	AGCGCCCCTA	AAGATATTTT
	AGACAACACTAC	CAAGATTGAA	TAATCATCAA	TCGCGGGGAT	TTCTATAAAA
2001	AGATAACCTT	CCGCAATTTT	TTTCTACTGT	TGATTTGCCA	ACTGACCAGA
	TCTATTGGAA	GGCGTTAAAG	AAAGATGACA	ACTAAACGGT	TGACTGGTCT
2051	TATTGATTGA	AGGATTAATT	TTCGAGGTTC	AGCAAGGTGA	TGCTTTAGAT
	ATAACTAACT	TCCTAATTAA	AAGCTCCAAG	TCGTTCCACT	ACGAAATCTA
2101	TTTTTCCTTTG	CTGCTGGCTC	TCAGCGCGGC	ACTGTTGCTG	GTGGTGTTAA
	AAAAGGAAAC	GACGACCGAG	AGTCGCGCCG	TGACAACGAC	CACCACAATT
2151	TACTGACCGT	CTAACCTCTG	TTTTATCTTC	TGCGGGTGGT	TCGTTTCGGTA
	ATGACTGGCA	GATTGGAGAC	AAAATAGAAG	ACGCCCACCA	AGCAAGCCAT
2201	TTTTTTAACGG	CGATGTTTTA	GGGCTATCAG	TTCGCGCATT	AAAGACTAAT
	AAAAATTGCC	GCTACAAAAT	CCCGATAGTC	AAGCGCGTAA	TTTCTGATTA
2251	AGCCATTCAA	AAATATTGTC	TGTGCCTCGT	ATTCTTACGC	TTTCAGGTCA
	TCGGTAAGTT	TTTATAACAG	ACACGGAGCA	TAAGAATGCG	AAAGTCCAGT
2301	GAAGGGTTCT	ATTTCTGTTG	GCCAGAATGT	CCCTTTTATT	ACTGGTTCGTG
	CTTCCCAAGA	TAAAGACAAC	CGGTCTTACA	GGGAAAATAA	TGACCAGCAC
2351	TAAGTGGTGA	ATCTGCCAAT	GTAAATAATC	CATTTTCAGAC	AATTGAGCGT
	ATTGACCACT	TAGACGGTTA	CATTTATTAG	GTAAAGTCTG	TTAACTCGCA
2401	CAAAATGTTG	GTATTTCTAT	GAGTGTTTTT	CCCGTTGCAA	TGGCTGGCGG
	GTTTTTACAAC	CATAAAGATA	CTCACAAAAA	GGGCAACGTT	ACCGACCGCC
2451	TAATATTGTT	TTAGATATAA	CCAGTAAGGC	CGATAGTTTG	AGTTCTTCTA
	ATTATAACAA	AATCTATATT	GGTCATTCCG	GCTATCAAAC	TCAAGAAGAT
2501	CTCAGGCAAG	TGATGTTATT	ACTAATCAAA	GAAGTATTGC	GACAACGGTT
	GAGTCCGTTT	ACTACAATAA	TGATTAGTTT	CTTCATAACG	CTGTTGCCAA
2551	AATTTGCGTG	ATGGTCAGAC	TCTTTTGCTC	GGTGGCCTCA	CTGATTACAA
	TTAAACGCAC	TACCAGTCTG	AGAAAACGAG	CCACCGGAGT	GACTAATGTT
2601	AAACACTTCT	CAAGATTCTG	GTGTGCCGTT	CCTGTCTAAA	ATCCCTTTAA
	TTTGTGAAGA	GTTCTAAGAC	CACACGGCAA	GGACAGATTT	TAGGGAAATT

28/39

2651	TCGGCCTCCT	GTTTAGCTCC	CGTTCTGATT	CTAACGAGGA	AAGCACGTTG
	AGCCGGAGGA	CAAATCGAGG	GCAAGACTAA	GATTGCTCCT	TTCGTGCAAC
2701	TACGTGCTCG	TCAAAGCAAC	CATAGTACGC	GCCCTGTAGC	GGCGCATTA
	ATGCACGAGC	AGTTTCGTTG	GTATCATGCG	CGGGACATCG	CCGCGTAATT
2751	GCGCGGCGGG	TGTGGTGGTT	ACGCGCAGCG	TGACCGCTAC	ACTTGCCAGC
	CGCGCCGCC	ACACCACCA	TGCGCGTCGC	ACTGGCGATG	TGAACGGTCG
2801	GCCCTAGCGC	CCGCTCCTTT	CGCTTTCTTC	CCTTCCTTTC	TCGCCACGTT
	CGGGATCGCG	GGCGAGGAAA	GCGAAAGAAG	GGAAGGAAAG	AGCGGTGCAA

BamHI

~~~~~

|      |             |            |            |            |            |
|------|-------------|------------|------------|------------|------------|
| 2851 | CTCCGGCTTT  | CCCCGTCAAG | CTCTAAATCG | GGGGATCCCT | TTAGGGTTCC |
|      | GAGGCCGAAA  | GGGGCAGTTC | GAGATTTAGC | CCCCTAGGGA | AATCCCAAGG |
| 2901 | GATTTAGTGC  | TTTACGGCAC | CTCGACCTCC | AAAACTTGA  | TTTGGGTGAT |
|      | CTAAATCACG  | AAATGCCGTG | GAGCTGGAGG | TTTTTGAACT | AAACCCACTA |
| 2951 | GGTTCACGTA  | GTGGGCCATC | GCCCTAATAG | ACGGTTTTTC | GCCCTTTGAC |
|      | CCAAGTGCAT  | CACCCGGTAG | CGGGATTATC | TGCCAAAAAG | CGGGAAACTG |
| 3001 | GTTGGAGTCC  | ACGTTCTTTA | ATAGTGGACT | CTTGTTCCAA | ACTGGAACAA |
|      | CAACCTCAGG  | TGCAAGAAAT | TATCACCTGA | GAACAAGGTT | TGACCTTGTT |
| 3051 | CACTCAACCC  | TATCTCGGTC | TATTCTTTTG | ATTTATAAGG | GATTTTGCCG |
|      | GTGAGTTGGG  | ATAGAGCCAG | ATAAGAAAAC | TAAATATTCC | CTAAAACGGC |
| 3101 | ATTTTCGGCCT | ATTGGTTAAA | AAATGAGCTG | ATTTAACAAA | AATTTAACGC |
|      | TAAAGCCGGA  | TAACCAATTT | TTTACTCGAC | TAAATTGTTT | TTAAATTGCG |
| 3151 | GAATTTTAAAC | AAAATATTTA | CGTTTACAAT | TTAAATATTT | GCTTATACAA |
|      | CTTAAAATTG  | TTTTATAATT | GCAAATGTTA | AATTTATAAA | CGAATATGTT |
| 3201 | TCTTCCTGTT  | TTTGGGGCTT | TTCTGATTAT | CAACCGGGGT | ACATATGATT |
|      | AGAAGGACAA  | AAACCCCGAA | AAGACTAATA | GTTGGCCCCA | TGTATACTAA |

ClaI

~~~~~

3251	GACATGCTAG	TTTTACGATT	ACCGTTCATC	GATTCTCTTG	TTTGCTCCAG
	CTGTACGATC	AAAATGCTAA	TGGCAAGTAG	CTAAGAGAAC	AAACGAGGTC
3301	ACTCTCAGGC	AATGACCTGA	TAGCCTTTTT	AGACCTCTCA	AAAATAGCTA
	TGAGAGTCCG	TTACTGGACT	ATCGGAAAAA	TCTGGAGAGT	TTTTATCGAT
3351	CCCTCTCCGG	CATGAATTTA	TCAGCTAGAA	CGGTTGAATA	TCATATTGAT
	GGGAGAGGCC	GTAATTAAAT	AGTCGATCTT	GCCAACTTAT	AGTATAACTA
3401	GGTGATTTGA	CTGTCTCCGG	CCTTTCTCAC	CCGTTTGAAT	CTTTACCTAC
	CCACTAAACT	GACAGAGGCC	GGAAAGAGTG	GGCAAACCTA	GAAATGGATG
3451	ACATTACTCA	GGCATTGCAT	TTAAAATATA	TGAGGGTTCT	AAAAATTTTT
	TGTAATGAGT	CCGTAACGTA	AATTTTATAT	ACTCCCAAGA	TTTTTAAAAA

29/39

3501	ATCCTTGCGT	TGAAATAAAG	GCTTCTCCCG	CAAAAGTATT	ACAGGGTCAT
	TAGGAACGCA	ACTTTATTTT	CGAAGAGGGC	GTTTTTCATA	TGTCCCAGTA
3551	AATGTTTTTG	GTACAACCGA	TTTAGCTTTA	TGCTCTGAGG	CTTTATTGCT
	TTACAAAAAC	CATGTTGGCT	AAATCGAAAT	ACGAGACTCC	GAAATAACGA
3601	TAATTTTGCT	AATTCTTTGC	CTTGCCTGTA	TGATTTATTG	GATGTTAACG
	ATTAAACGA	TTAAGAAACG	GAACGGACAT	ACTAAATAAC	CTACAATTGC
3651	CTACTACTAT	TAGTAGAATT	GATGCCACCT	TTTCAGCTCG	CGCCCCAAAT
	GATGATGATA	ATCATCTTAA	CTACGGTGGA	AAAGTCGAGC	GCGGGGTTTA
3701	GAAAATATAG	CTAAACAGGT	TATTGACCAT	TTGCGAAATG	TATCTAATGG
	CTTTTATATC	GATTTGTCCA	ATAACTGGTA	AACGCTTTAC	ATAGATTACC
3751	TCAAACATAA	TCTACTCGTT	CGCAGAATTG	GGAATCAACT	GTTACATGGA
	AGTTTGATTT	AGATGAGCAA	GCGTCTTAAC	CCTTAGTTGA	CAATGTACCT
3801	ATGAAACTTC	CAGACACCGT	ACTTTAGTTG	CATATTTAAA	ACATGTTGAG
	TACTTTGAAG	GTCTGTGGCA	TGAAATCAAC	GTATAAATTT	TGTACAACCT
3851	CTACAGCACC	AGATCCAGCA	ATTAAGCTCT	AAGCCATCCG	CAAAAATGAC
	GATGTCGTGG	TCTAGGTCGT	TAATTCGAGA	TTCGGTAGGC	GTTTTTACTG
3901	CTCTTATCAA	AAGGAGCAAT	TAAAGGTACT	CTCTAATCCT	GACCTGTTGG
	GAGAATAGTT	TTCCTCGTTA	ATTTCCATGA	GAGATTAGGA	CTGGACAACC
3951	AGTTTGCTTC	CGGTCTGGTT	CGCTTTGAAG	CTCGAATTAA	AACGCGATAT
	TCAAACGAAG	GCCAGACCAA	GCGAAACTTC	GAGCTTAATT	TTGCGCTATA
4001	TTGAAGTCTT	TCGGGCTTCC	TCTTAATCTT	TTTGATGCAA	TCCGCTTTGC
	AACTTCAGAA	AGCCCGAAGG	AGAATTAGAA	AAACTACGTT	AGGCGAAACG
4051	TTCTGACTAT	AATAGTCAGG	GTAAAGACCT	GATTTTTTGAT	TTATGGTCAT
	AAGACTGATA	TTATCAGTCC	CATTTCTGGA	CTAAAACTA	AATACCAGTA
4101	TCTCGTTTTT	TGAACTGTTT	AAAGCATTTG	AGGGGGATTG	AATGAATATT
	AGAGCAAAAG	ACTTGACAAA	TTTCGTAAAC	TCCCCCTAAG	TTACTTATAA
4151	TATGACGATT	CCGCAGTATT	GGACGCTATC	CAGTCTAAAC	ATTTTACTAT
	ATACTGCTAA	GGCGTCATAA	CCTGCGATAG	GTCAGATTTG	TAAAATGATA
4201	TACCCCCCTC	GGCAAAACTT	CTTTTGCAAA	AGCCTCTCGC	TATTTTTTGT
	ATGGGGGAGA	CCGTTTTGAA	GAAAACGTTT	TCGGAGAGCG	ATAAAAACAA
4251	TTTATCGTCG	TCTGGTAAAC	GAGGGTTATG	ATAGTGTTGC	TCTTACTATG
	AAATAGCAGC	AGACCATTTG	CTCCCAATAC	TATCACAACG	AGAATGATAC
4301	CCTCGTAATT	CCTTTTGCGG	TTATGTATCT	GCATTAGTTG	AATGTGGTAT
	GGAGCATTA	GGAAAACCGC	AATACATAGA	CGTAATCAAC	TTACACCATA
4351	TCCTAAATCT	CAACTGATGA	ATCTTTCTAC	CTGTAATAAT	GTTGTTCCGT
	AGGATTTAGA	GTTGACTACT	TAGAAAGATG	GACATTATTA	CAACAAGGCA

30/39

4401	TAGTTCGTTT	TATTAACGTA	GATTTTTCTT	CCCAACGTCC	TGACTGGTAT
	ATCAAGCAAA	ATAATTGCAT	CTAAAAAGAA	GGGTTGCAGG	ACTGACCATA
4451	AATGAGCCAG	TTCTTAAAAT	CGCATAAGGT	AATTCACAAT	GATTAAAGTT
	TTACTCGGTC	AAGAATTTTA	GCGTATTCCA	TTAAGTGTTA	CTAATTTCAA
4501	GAAATTAAAC	CATCTCAAGC	GCAATTCACT	ACCCGTTCTG	GTGTTTCTCG
	CTTTAATTTG	GTAGAGTTTCG	CGTTAAGTGA	TGGGCAAGAC	CACAAAGAGC
4551	TCAGGGCAAG	CCTTATTCAC	TGAATGAGCA	GCTTTGTTCAC	GTTGATTTGG
	AGTCCCGTTC	GGAATAAGTG	ACTTACTCGT	CGAAACAATG	CAACTAAACC
4601	GTAATGAATA	TCCGGTGCTT	GTCAAGATTA	CTCTTGATGA	AGGTCAGCCA
	CATTACTTAT	AGGCCACGAA	CAGTTCTAAT	GAGAACTACT	TCCAGTCGGT
4651	GCCTATGCGC	CTGGTCTGTA	CACCGTGCAT	CTGTCCTCGT	TCAAAGTTGG
	CGGATACGCG	GACCAGACAT	GTGGCACGTA	GACAGGAGCA	AGTTTCAACC
4701	TCAGTTCGGT	TCTCTTATGA	TTGACCGTCT	GCGCCTCGTT	CCGGCTAAGT
	AGTCAAGCCA	AGAGAATACT	AAC TGGCAGA	CGCGGAGCAA	GGCCGATTCA
4751	AACATGGAGC	AGGTCGCGGA	TTTCGACACA	ATTTATCAGG	CGATGATACA
	TTGTACCTCG	TCCAGCGCCT	AAAGCTGTGT	TAAATAGTCC	GCTACTATGT
4801	AATCTCCGTT	GTACTTTGTT	TCGCGCTTGG	TATAATCGCT	GGGGGTCAAA
	TTAGAGGCAA	CATGAAACAA	AGCGCGAACC	ATATTAGCGA	CCCCCAGTTT
4851	GATGAGTGTT	TTAGTGTATT	CTTTCGCCTC	TTTCGTTTTA	GGTTGGTGCC
	CTACTCACAA	AATCACATAA	GAAAGCGGAG	AAAGCAAAAT	CCAACCACGG
4901	TTCGTAGTGG	CATTACGTAT	TTTACCCGTT	TAATGGAAAC	TTCTTCATGC
	AAGCATCACC	GTAATGCATA	AAATGGGCAA	ATTACCTTTG	AAGGAGTACG
4951	GTAAGTCTTT	AGTCCTCAAA	GCCTCCGTAG	CCGTTGCTAC	CCTCGTTCCG
	CATTTCAGAAA	TCAGGAGTTT	CGGAGGCATC	GGCAACGATG	GGAGCAAGGC
5001	ATGCTGTCTT	TCGCTGCTGA	GGGTGACGAT	CCCGCAAAAG	CGGCCTTTGA
	TACGACAGAA	AGCGACGACT	CCCCTGCTA	GGGCGTTTTT	GCCGGAAACT
5051	CTCCCTGCAA	GCCTCAGCGA	CCGAATATAT	CGGTTATGCG	TGGGCGATGG
	GAGGGACGTT	CGGAGTCGCT	GGCTTATATA	GCCAATACGC	ACCCGCTACC
5101	TTGTTGTCAT	TGTCGGCGCA	ACTATCGGTA	TCAAGCTGTT	TAAGAAATTC
	AACAACAGTA	ACAGCCGCGT	TGATAGCCAT	AGTTCGACAA	ATTCTTTAAG
5151	ACCTCGAAAG	CAAGCTGATA	AAGGAGGTTT	CTCGATCGAG	ACGTTGGGTG
	TGGAGCTTTC	GTTTCGACTAT	TTCTTCCAAA	GAGCTAGCTC	TGCAACCCAC
5201	AGGTTCCAAC	TTTCACCATA	ATGAAATAAG	ATCACTACCG	GGCGTATTTT
	TCCAAGGTTG	AAAGTGGTAT	TACTTTTATTC	TAGTGATGGC	CCGCATAAAA
5251	TTGAGTTATC	GAGATTTTCA	GGAGCTAAGG	AAGCTAAAAT	GGAGAAAAAA
	AACTCAATAG	CTCTAAAAGT	CCTCGATTCC	TTCGATTTTA	CCTCTTTTTT

31/39

5301	ATCACTGGAT TAGTGACCTA	ATACCACCGT TATGGTGGCA	TGATATATCC ACTATATAGG	CAATGGCATC GTTACCGTAG	GTAAAGAACA CATTTCTTGT
5351	TTTTGAGGCA AAACTCCGT	TTTCAGTCAG AAAGTCAGTC	TTGCTCAATG AACGAGTTAC	TACCTATAAC ATGGATATTG	CAGACCGTTC GTCTGGCAAG
5401	AGCTGGATAT TCGACCTATA	TACGGCCTTT ATGCCGGAAA	TTAAAGACCG AATTTCTGGC	TAAAGAAAAA ATTTCTTTTT	TAAGCACAAG ATTCGTGTTC
5451	TTTTATCCGG AAAATAGGCC	CCTTTATTCA GGAAATAAGT	CATTCTTGCC GTAAGAACGG	CGCCTGATGA GCGGACTACT	ATGCTCATCC TACGAGTAGG
5501	GGAGTTCCGT CCTCAAGGCA	ATGGCAATGA TACCGTTACT	AAGACGGTGA TTCTGCCACT	GCTGGTGATA CGACCACTAT	TGGGATAGTG ACCCTATCAC
5551	TTCACCCTTG AAGTGGGAAC	TTACACCGTT AATGTGGCAA	TTCCATGAGC AAGGTACTCG	AAACTGAAAC TTTGACTTTG	GTTTTTCATCG CAAAAGTAGC
5601	CTCTGGAGTG GAGACCTCAC	AATACCACGA TTATGGTGCT	CGATTTCCGG GCTAAAGGCC	CAGTTTCTAC GTCAAAGATG	ACATATATTC TGTATATAAG
5651	GCAAGATGTG CGTTCTACAC	GCGTGTTACG CGCACAATGC	GTGAAAACCT CACTTTTGGA	GGCCTATTTT CCGGATAAAG	CCTAAAGGGT GGATTTCCCA
5701	TTATTGAGAA AATAACTCTT	TATGTTTTTC ATACAAAAG	GTCTCAGCCA CAGAGTCGGT	ATCCCTGGGT TAGGGACCCA	GAGTTTCACC CTCAAAGTGG
5751	AGTTTTGATT TCAAACTAA	TAAACGTAGC ATTTGCATCG	CAATATGGAC GTTATACCTG	AAC TTCTTCG TTGAAGAAGC	CCCCCGTTTT GGGGGCAAAA
5801	CACTATGGGC GTGATACCCG	AAATATTATA TTTATAATAT	CGCAAGGCGA GCGTTCGCT	CAAGGTGCTG GTTCCACGAC	ATGCCGCTGG TACGGCGACC
5851	CGATTCAGGT GCTAAGTCCA	TCATCATGCC AGTAGTACGG	GTTTGTGATG CAAACACTAC	GCTTCCATGT CGAAGGTACA	CGGCAGAATG GCCGTCTTAC
5901	CTTAATGAAT GAATTACTTA	TACAACAGTA ATGTTGTCAT	CTGCGATGAG GACGCTACTC	TGGCAGGGCG ACCGTCCCGC	GGGCGTAATT CCCGCATTA
5951	TTTTTAAGGC AAAAATTCCG	AGTTATTGGT TCAATAACCA	GCCCTTAAAC CGGGAATTTG	GCCTGGTGCT CGGACCACGA	AGCCTGAGGC TCGGACTCCG
6001	CAGTTTGCTC GTCAAACGAG	AGGCTCTCCC TCCGAGAGGG	CGTGGAGGTA GCACCTCCAT	ATAATTGCTC TATTAACGAG	GACCGATAAA CTGGCTATTT
6051	AGCGGCTTCC TCGCCGAAGG	TGACAGGAGG ACTGTCCTCC	CCGTTTTGTT GGCAAAACAA	TTGCAGCCCA AACGTCGGGT	CCTCAACGCA GGAGTTGCGT
6101	ATTAATGTGA TAATTACACT	GTTAGCTCAC CAATCGAGTG	TCATTAGGCA AGTAATCCGT	CCCCAGGCTT GGGGTCCGAA	TACACTTTAT ATGTGAAATA
6151	GCTTCCGGCT CGAAGGCCGA	CGTATGTTGT GCATACAACA	GTGGAATTGT CACCTTAACA	GAGCGGATAA CTCGCCTATT	CAATTTTACA GTTAAAGTGT

32/39

6201	CAGGAAACAG	CTATGACCAT	GATTACGAAT	TTCTAGATAA	CGAGGGCAAA
	GTCCTTTGTC	GATACTGGTA	CTAATGCTTA	AAGATCTATT	GCTCCCGTTT
6251	AAATGAAAAA	GACAGCTATC	GCGATTGCAG	TGGCACTGGC	TGGTTTCGCT
	TTTACTTTTT	CTGTCGATAG	CGCTAACGTC	ACCGTGACCG	ACCAAAGCGA
6301	ACCGTAGCGC	AGGCCGACTA	CAAAGATGTC	GACTGTATTG	TTTATCATGC
	TGGCATCGCG	TCCGGCTGAT	GTTTCTACAG	CTGACATAAC	AAATAGTACG

BamHI EcoRI

~~~~~

|      |            |            |            |            |            |
|------|------------|------------|------------|------------|------------|
| 6351 | TCATTATCTT | GTTGCTAAGT | GTGGTGGTGG | AGGATCCGAA | TTCAATGCTG |
|      | AGTAATAGAA | CAACGATTCA | CACCACCACC | TCCTAGGCTT | AAGTTACGAC |
| 6401 | GCGGCGGCTC | TGGTGGTGGT | TCTGGTGGCG | GCTCTGAGGG | TGGTGGCTCT |
|      | CGCCGCCGAG | ACCACCACCA | AGACCACCGC | CGAGACTCCC | ACCACCGAGA |
| 6451 | GAGGGTGGCG | GTTCTGAGGG | TGGCGGCTCT | GAGGGAGGCG | GTTCCGGTGG |
|      | CTCCCACCGC | CAAGACTCCC | ACCGCCGAGA | CTCCCTCCGC | CAAGGCCACC |
| 6501 | TGGCTCTGGT | TCCGGTGATT | TTGATTATGA | AAAGATGGCA | AACGCTAATA |
|      | ACCGAGACCA | AGGCCACTAA | AACTAATACT | TTTCTACCGT | TTGCGATTAT |
| 6551 | AGGGGGCTAT | GACCGAAAAT | GCCGATGAAA | ACGCGCTACA | GTCTGACGCT |
|      | TCCCCCGATA | CTGGCTTTTA | CGGCTACTTT | TGCGCGATGT | CAGACTGCGA |

ClaI

~~~~~

6601	AAAGGCAAAC	TTGATTCTGT	CGCTACTGAT	TACGGTGCTG	CTATCGATGG
	TTTCCGTTTG	AACTAAGACA	GCGATGACTA	ATGCCACGAC	GATAGCTACC
6651	TTTCATTGGT	GACGTTTCCG	GCCTTGCTAA	TGGTAATGGT	GCTACTGGTG
	AAAGTAACCA	CTGCAAAGGC	CGGAACGATT	ACCATTACCA	CGATGACCAC
6701	ATTTTGCTGG	CTCTAATTCC	CAAATGGCTC	AAGTCGGTGA	CGGTGATAAT
	TAAAACGACC	GAGATTAAGG	GT'TTACCGAG	TTCAGCCACT	GCCACTATTA
6751	TCACCTTTAA	TGAATAATTT	CCGTCAATAT	TTACCTTCCC	TCCCTCAATC
	AGTGGAATTT	ACTTATTAAA	GGCAGTTATA	AATGGAAGGG	AGGGAGTTAG
6801	GGTTGAATGT	CGCCCTTTTG	TCTTTGGCGC	TGGTAAACCA	TATGAATTTT
	CCAACCTTACA	GCGGGAAAAC	AGAAACCGCG	ACCATTTGGT	ATACTTAAAA
6851	CTATTGATTG	TGACAAAATA	AACTTATTCC	GTGGTGTCTT	TGCGTTTCTT
	GATAACTAAC	ACTGTTTTAT	TTGAATAAGG	CACCACAGAA	ACGCAAAGAA
6901	TTATATGTTG	CCACCTTTAT	GTATGTATTT	TCTACGTTTG	CTAACATACT
	AATATACAAC	GGTGGAATAA	CATACATAAA	AGATGCAAAC	GATTGTATGA

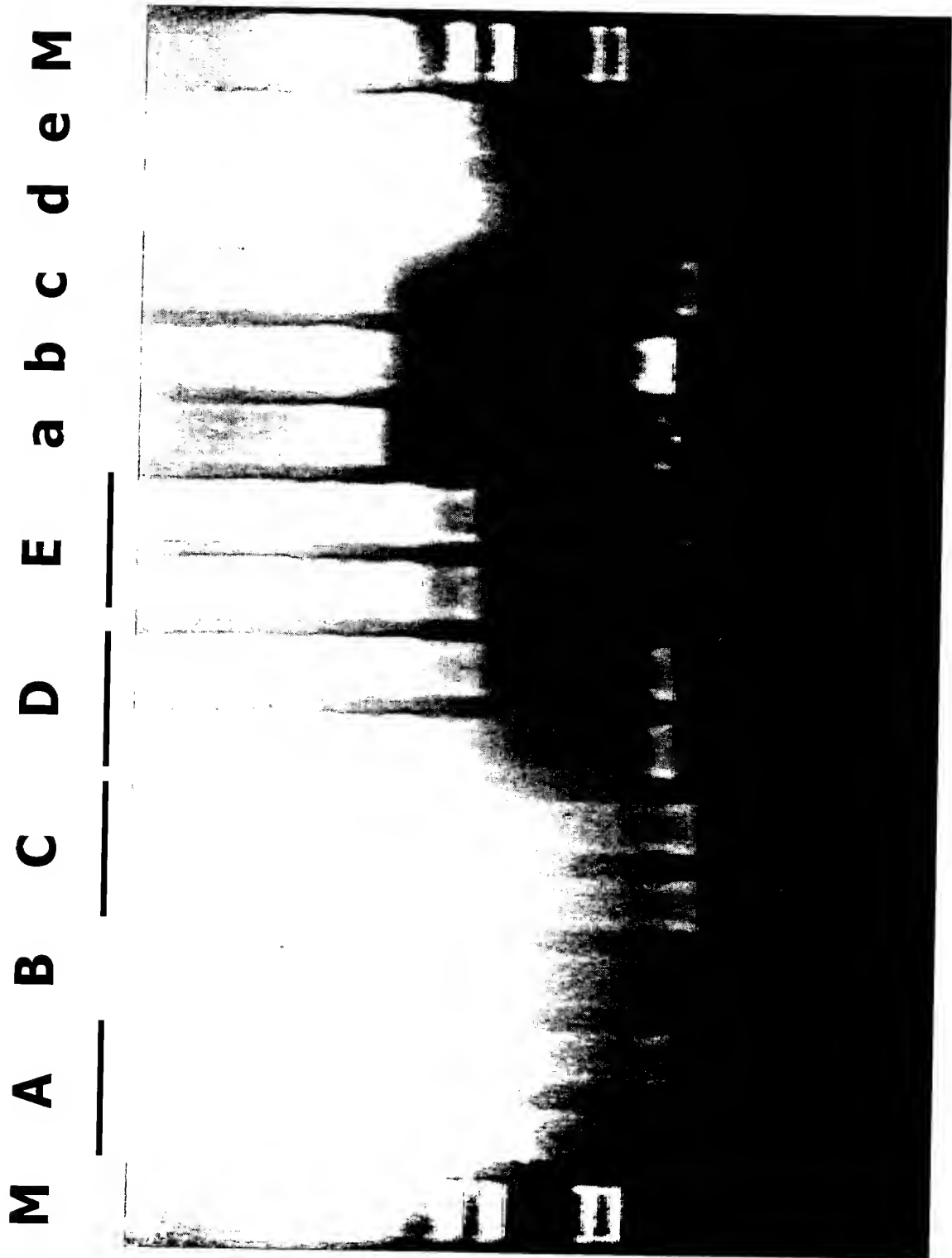
HindIII

33/39

6951 GCGTAATAAG GAGTCTTGAT A
CGCATTATTC CTCAGAACTA T

34/39

Figure 5



35/39

Figure 6
M SIP Polypophage transductants transf.
CO-

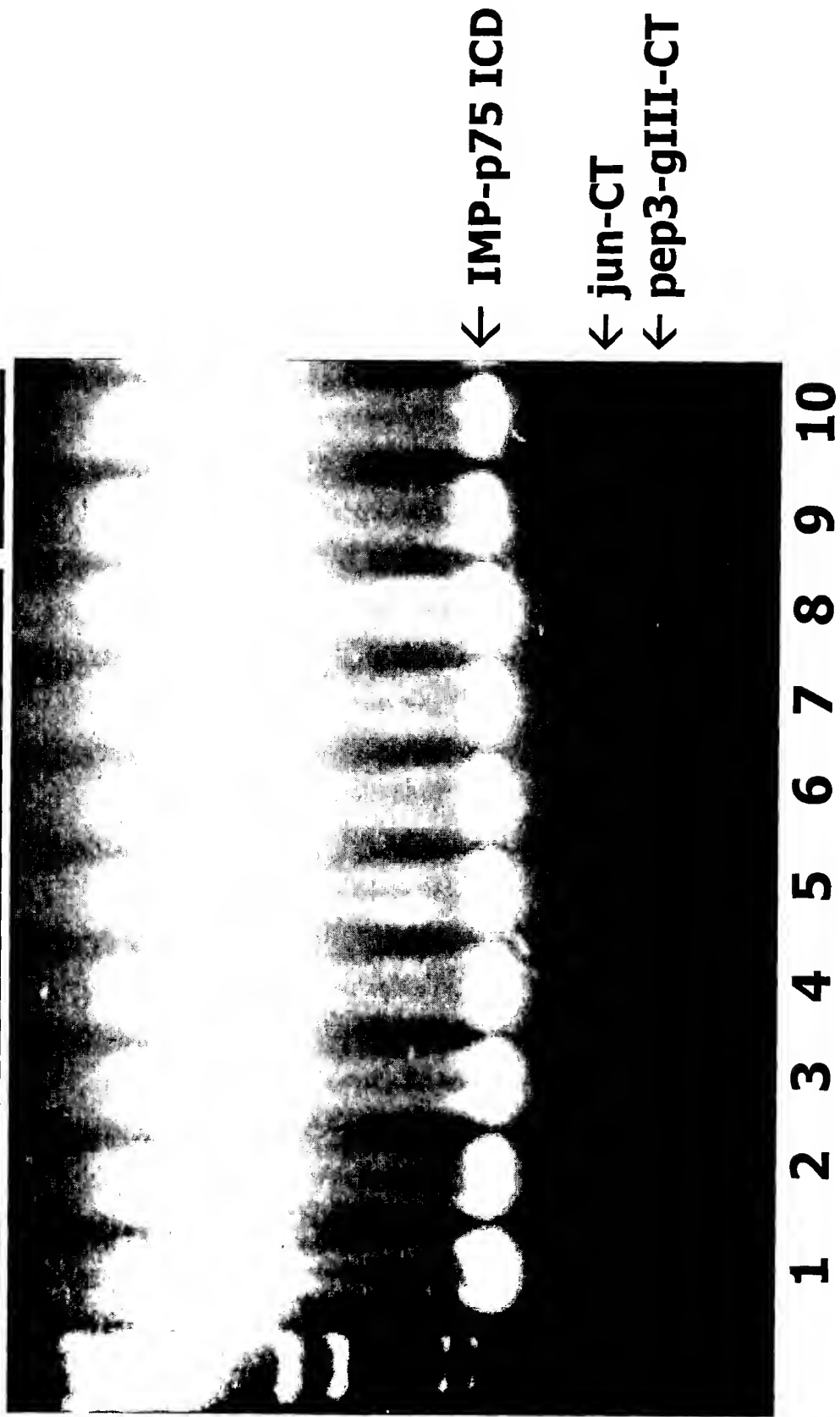


Figure 7

dilution factor		transductants (t.u./ml)*
pep3/p75ICD	jun/p75ICD	
1	pos. control -	6×10^5
-	neg. control 1	0
1	10^2	1.2×10^4
1	10^3	8.6×10^2
1	10^4	1.2×10^2
1	10^5	12 [#]
1	10^6	1.2 [#]
1	10^7	0.12 [#]

37/39

Figure 8

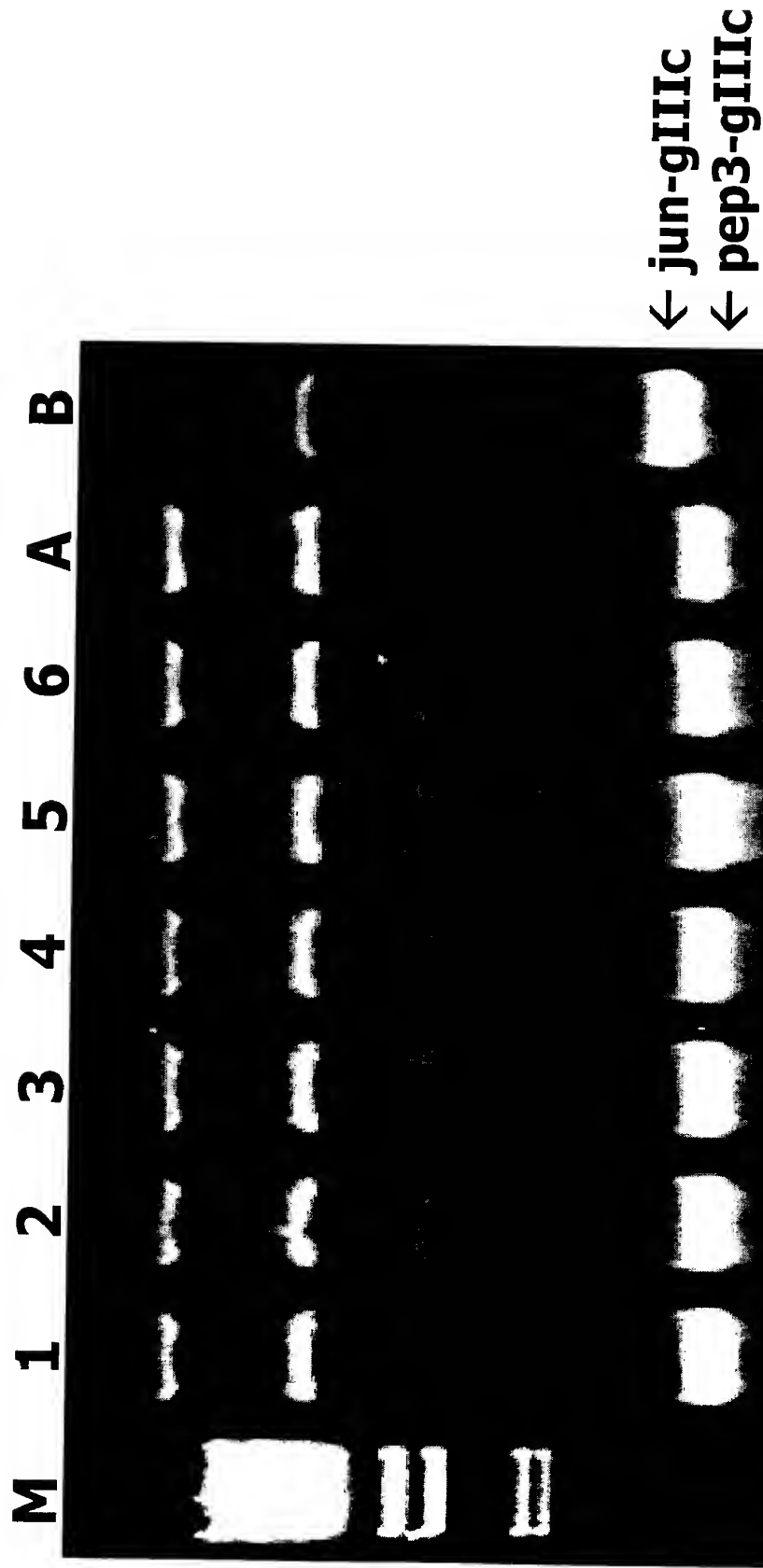
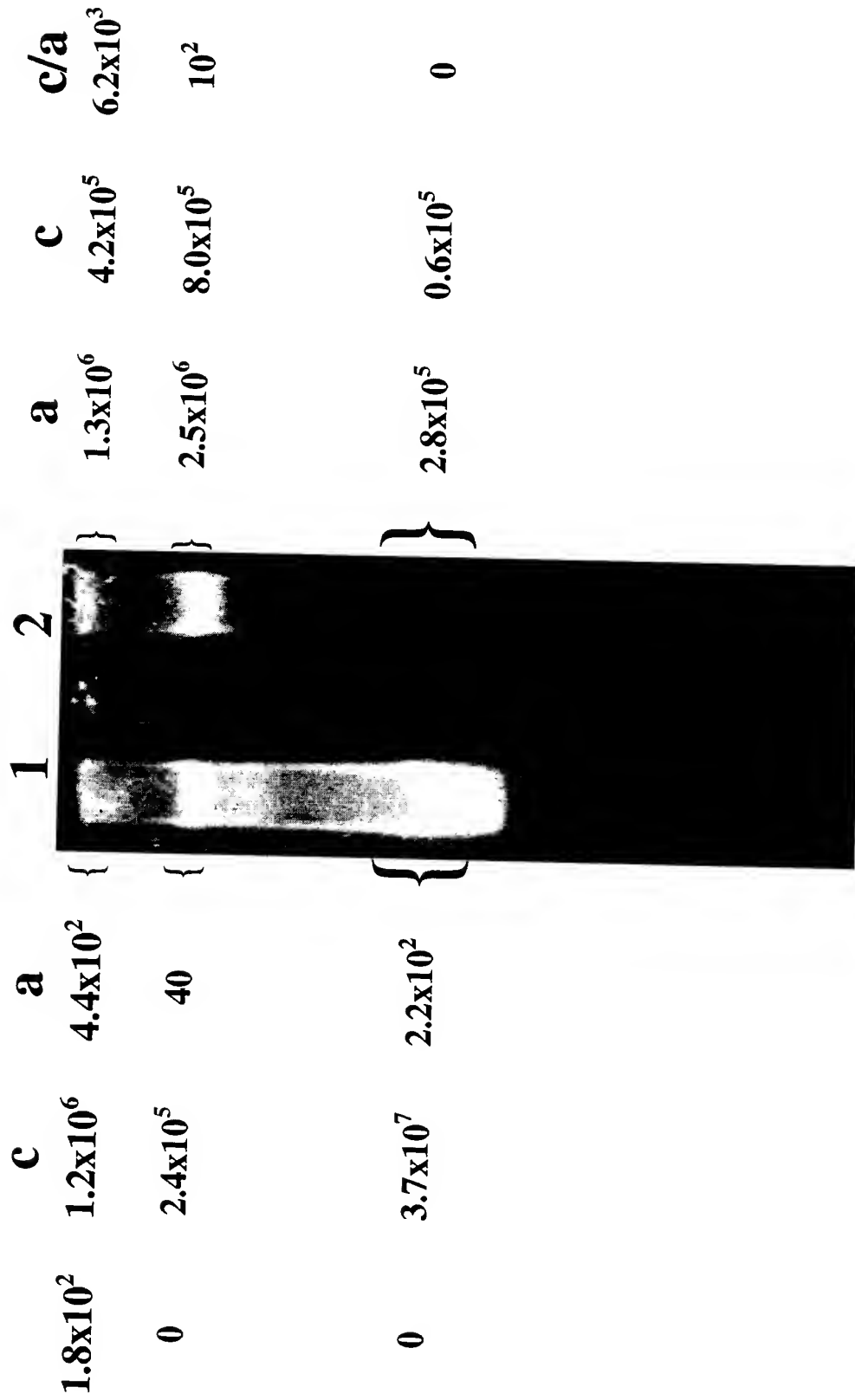


Figure 9



M a/b c/d a/b c/d a/b c/d a d c b M

